Workshop Discussion Notes: Employment

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

This document was produced based on notes taken during the Employment workshop 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. All workshop participants received workshop materials prior to the event to spark discussion. The primer can be found at: http://www.datacivilrights.org/pubs/2014-1030/Employment.pdf

Overview

This session approached the question of Big Data's growing influence on employment, with a specific focus on hiring practices. The group thought through the applicability of a "due process" framework to the question of data and employment, particularly to address the problem of how job candidate profiles can be marked with inaccurate information or inaccurate inferences, and what redress or remedies are available to those who are subsequently excluded from job opportunities. The elements of due process are able to inform the actions taken to rectify this, yet it's necessary to understand due processes' own potential downfalls—the group worked together to assemble a large spectrum of possible problems with applying the concept to data practices. Finally, the group brainstormed on how big data techniques compare and interact with other processes of hiring—how do we make comparative claims between hiring techniques? How do we establish what works and what does not? And for whom? And how do these questions apply to the increasing variety of non-traditional employment scenarios that eschew the familiar application-interview-acceptance model?

Themes and Discussion Topics

How applicable is a "due process" framework to the evaluation of Big Data's effect on the sphere of employment? In many cases of data use, the best case scenario is determining what can fairly and accurately be done with "good" or correct data. Such a focus on good uses of good data, however, overlooks the very real possibility of having to work from "bad" or incorrect data. In certain contexts to which the Fair Credit Reporting Act (FCRA) applies, such as in decisions made about individuals credit scores, organizations like FICO are obliged to maintain accurate records on individuals. However, outside of these contexts, all sorts of scoring mechanisms can be generated from the data gathered on people, and the inferences attributed to them, and there is no similar obligation for data to be accurate. For instance, there are databases where people's data profiles reside under the auspices of opinion, rather than fact, but if this data is inaccurate or inferences are made about it that are harmful, and these data or inferences are used in algorithmic scoring mechanisms, they can produce harmful results, with no remedy available to the individuals affected them. This is where a due process framework has the chance to produce an intervention—by providing a mechanism for evaluating why certain algorithmic decisions get made (even if they are not initially legible to human analysis).

One complicating factor is the use of proxies in data work. Though, there is some confusion over what counts as a data proxy that stems from the variety of conditions indicated by the term. Proxies can be close correlates to protected status (and thus a numerical value). Proxies might refer to the relationship between data categories—isolating one dimension causes another dimension to become less predictive. For example, does golf digest operate as a different predictor of income between black and white populations of subjects? It is productive to think of proxies as a spectrum rather than a simple binary.

Applying a due process framework to employment implies examining the data produced by hiring decisions. This is partially because due process is about addressing adjudication, and hiring decisions are where adjudication happens in the commercial setting. However, hiring decisions are just the end of a much longer chain of procedures. This is a struggle common to many laws meant to address civil rights violations—focusing on final decisions alone can mask earlier discriminatory structures. For instance, the Equal Employment Opportunity Commission (EEOC) publishes no advice on the issue of how past hiring biases might create a discriminatory set of data for subsequent decisions made by algorithm, although case law can provide important precedents for guidance in this area, and there are well-established guidelines for identifying bias in employment practices. They do have guidelines related to structurally similar cases, but applying them takes some creative application. The example cited here was an older case regarding firefighters in Warren, a suburb of Detroit. The Fire Department advertised in local papers, unfairly limiting the “search” to white populations, and therefore had only white employees, and they were subsequently held accountable by the EEOC for discriminatory hiring practices. In the end, due process is not a silver bullet to questions of data-influenced hiring practices. It doesn’t address earlier structures of discrimination, and it doesn’t currently apply to private (rather than public) employees.

In applying data techniques to discriminatory algorithms in hiring and employment contexts, what are the potential problems that can create discriminatory conditions? The group agreed on a diverse list of potential pitfalls:

- The data used might be inaccurate/erroneous
- The decision might be based on the “wrong” dataset (For example, if all the people who have ever worked in a particular capacity were women (due to historical prejudices) then a new algorithm trained on that population risks inadvertently selecting for that already imbalanced criteria.
- The data might be irrelevant/unrelated to actual job performance
- There might be a culturally unaware metric implemented
    - Although not related specifically to hiring, but rather, to broader issues of discrimination that *could* apply to employment, the group discussed the example of Ms. Elaine Yellow Horse, whose Google+ account was suspended because it appeared that her minority name was not real to the Google+ algorithm, in violation of their name policy (the policy has since been modified). She protested that the mail-order coffee business she helps to run suffered as a result of her inability to use/sign-in to her Google+ account for business meetings, for example.
- Algorithmic decisions could be made malicious by design
- Algorithmic decisions could be “bad” through accident or ineptitude

- A lack of transparency/opacity
- Relevant data might be legally inadmissible (e.g. LGBT status)

In addition to these dimensions that the group broadly agreed on, there was further discussion of edge cases and less clear possibilities. The intentional conflation of subjective and objective data can be a problem, in both directions. Objective data can be treated as subjective to mask inconvenient facts, and subjective data can be treated as objective to produce unjustified inferences. It was also pointed out that algorithms with truly "neutral" (without intentional bias) design can still have disparate impacts on various populations. This leads to a still unclear domain of definitions: how do you define fairness to a machine learning algorithm? How do you consistently specify which criteria are "valid"? Finally, does the very real possibility tell us something about bad actors, or does it just reflect the already operational cultural biases that even good faith actors are unwittingly implementing? Is machine learning always fated to be as discriminatory as culture at large?

After discussing the potential challenges to implementing a system based on data analysis, the group discussed how these techniques compare and relate to the larger sphere of processes that structure employment as is. In many cases there are already challenges and inequalities at work before any data techniques are introduced. In these instances, data analysis can actually reveal problems and create new opportunities. For instance, some data findings indicate college degrees to be less valuable than is widely assumed. There are many reasons for this—possession of a college degree is a convenient shorthand for asking about a collection of disparate skills. In general, it is hard to code skills in ways that translate well across employment contexts. For example, many military vets who have a considerable number of skills that are useful but do not fit the assumed model of many employers. Lack of an MBA might unfairly be interpreted as lack of managerial skills. There are examples of employers beginning to proactively use data analysis to compensate for these disparities among differently-qualified candidates. One participant described a program being implemented at Google to do just this.

Data analysis is affecting how companies measure employee performance, but this also feeds back into the hiring process. An uneven comparison can result—between those you hired (on whom employers have lots of data) and those you did not (on whom employers have almost none). This observation led the group to a re-contextualization: companies aren't going to be able to use data to eliminate subtle decision making about people. There will always be "soft" things at work in interpersonal relationships. The larger question is whether those types of decisions are being made better or worse by the introduction of data, and for whom. Employment with McDonalds or Radio Shack, for instance, is dependent on an online personality quiz meant to reveal a history of mental illness. Without awareness on the part of the applicant of the function of this test, certain answers can inadvertently disqualify them. However, this can amount to an unlawful pre-employment medical exam and is the subject of an ongoing lawsuit.

Areas for Further Exploration

It is apparent that hiring practices can be significantly affected by the adoption of data technologies, and that these effects are not inherently positive or negative. After discussing the variety of impacts such technologies can have on hiring, the group began to acknowledge that the field of influence extended well beyond the traditional application-interview-acceptance model. There are job opportunities that are never applied for, but are filled by mediating vendors that can create custom “applicant” pools from standing databases. Even in cases where algorithms and data technologies can have a positive impact (on employee effectiveness, or fairness of hiring practices) the fact that such technologies *also* introduce an opaque process that cannot be rationalized by a human mind might be reason enough to avoid them. Therefore the *clarity* of this process, in addition to the *equity* of it, might become a policy priority for implementers and regulators.