

# 5 Questions on Data & Justice with Cathy O'Neil



Catherine D'Ignazio [Follow](#)

Nov 26, 2017 · 8 min read



Cathy O'Neil. Image from [mathbabe.org](http://mathbabe.org).

Cathy O'Neil is a data scientist, [blogger](#), contributing columnist at [Bloomberg](#), and author of the recent book [Weapons of Math Destruction](#). She is an eloquent and urgent voice on the inequality present in algorithms, particularly those that use big data to sort people into categories (like “Good teacher/Bad teacher”, “Consider for hiring/Don’t consider for hiring”, “High risk for reoffending/Low risk for reoffending”). These kinds of decisions have real, material consequences for people’s lives. And while the automated systems that drive them may have the appearance of neutrality, they are often using inputs like zip code or last name that are proxies for race and ethnicity. The

risk, for individuals and society more broadly, is that computational decision-making risks re-encoding structural inequalities around race, gender, ethnicity and socioeconomic status. “Garbage in, Garbage out,” as they say in programming. Or, put more bluntly, “Racism in, Racism out.”

I caught up with O'Neil at the inaugural [Data for Black Lives conference](#) at the MIT Media Lab in Cambridge, MA. O'Neil spoke on [the keynote panel](#) about algorithmic predictions, ethics training for data scientists, and a “Data Bill of Rights” that would give individuals insight into how they were being sorted and rated by automated systems. Below is a transcript of our conversation.

## **You characterize some algorithms as “Weapons of Math Destruction.” Who is most vulnerable to being harmed by their use?**

In some sense, this is a tautological question because I only call it a weapon of math destruction *because* it is destructive. That means it makes unlucky people unluckier. By people being unlucky I mean that they face a bias that is due to a characteristic beyond their control — like racist, sexist, classist, or religious bias. There was an example in my book where Kyle Beam was prevented from getting a job because he had a particular mental health status. So, this could include people with disabilities.

Now, there are algorithms called price discrimination algorithms that punish people that aren't typically unlucky. The price discrimination concept is to figure out how much someone is willing to pay and charge them that. Now these can be problematic and I talk about that in the book in the insurance chapter because these can effectively discriminate against people who are poor. So it's like charging people more for their lack of understanding of the world. A tax on ignorance, effectively. But there's lots of examples of price discrimination that basically like, “Oh, you're rich, we are going to charge you more.” There was [a famous case](#) of Orbitz charging people with Macs more than they charge people with PCs. I don't really feel bad about that. Charging rich people more doesn't really bother me.

One of the most interesting aspects of this is that what we decide is troubling is a moral question. It has nothing to do with technology. Technology just makes it easier to hide these immoral decisions in the guise of something that seems objective.

## **What is (or what could be) the role of white men in data justice?**

I just wrote a piece this week for Bloomberg View about Ray Dalio's algorithm at Bridgewater (*Interviewer's note: Bridgewater is the world's biggest hedge fund and it was founded by Dalio*). The main point of the piece is that there is such a thing as an "alpha male bias". For example, when you have people judging other people's ideas and scoring those ideas on various criteria like, "Is it original? Is it important? Is it going to make us money? Is it convincing? How long will it last?"

We think of that as an objective scoring system. But actually the same idea posed by a young, beta male versus an older, alpha male is going to be scored very differently. The reason this matters is because there was a sexual harassment issue at Bridgewater. As the Wall Street Journal reported it, part of the way that it was sorted out is that they looked at the "believability scores" of the guy and the woman. The guy was the co-chief investment officer and the woman was unnamed and junior and reported to the guy. At Bridgewater they depend very heavily on these metrics and one of them is called "believability". And it's a bogus system but they believe in it so that's what I was writing about.

So this particular Bloomberg article was very popular with the Bloomberg readership. So popular that I was surprised. I write about racial bias, gender bias, and classism all the time but this is the one that became popular. But then I realized that of course it's because white males read Bloomberg and they finally get it and can relate to something in their lives. Because after all, everyone hates their boss.

So my new "in" with the white male audience is this alpha male bias. This makes it clear to them that these systems are not objective and I can say, "Hey, you know what else is not objective?"

But it's hard with white guys in fancy, corporate fields. So, my real answer to the question is that we have a long road ahead of us. Some guys are obviously already on top of it and others will need help. But relatable examples are a start.

**You have been advocating for a Bill of Rights around data — what would that entail and who would it be for?**

It's for everyone and it applies at all times. That's the important part. There are a couple algorithms I write about in the book. For example, the teacher value-added model which is assessing teachers and, in the worst case scenario, firing them when they get bad

scores. There's the job application scoring system that I talk about with Kyle. If you don't get a high enough score you don't even get an interview. So in the worst case scenario there you are denied access to employment. And the third is the recidivism risk score where the worst case is that you are sent to prison for longer or denied bail or parole.

These are important things and it seems insane to me that people's scores are secret to them. They don't understand the system. They can't understand the formula. The systems are secret and there is no mechanism to appeal. Recently in Houston, TX, there was a judge that ruled that six teachers who had been fired based on their value-added model score had been denied due process. It is my fantasy that this same argument would hold in all those other situations, too.

To me, a Bill of Rights means that you cannot deny people important opportunities or their liberty based on secret systems with no method of appeal. We have civil rights laws. This is a civil rights issue.

**Since you have published the book, do you see a growing conversation around data & justice? Who are some of the people and communities that you have encountered that give you hope?**

I'm seeing larger group of people getting involved who are activists, lawyers, even government officials, at the local and federal level. It's been great.

I wrote an op-ed for the New York Times this week that a lot of people hated about how the academic "machine" has to do more. People interpreted it in a way I didn't mean them to. They thought I was saying that academics aren't doing anything. That's not what I meant. I intended to say that the machine of academia itself isn't doing anything — so, administrators if you will. They allow small non-STEM fields to think about critique of algorithms but then starve those fields for tenure lines and institutional resources while they build the computer science department a new building and give them extra faculty lines. Then, they accept post-docs funded by Google and Facebook and they create data science masters programs that are cash cows that don't teach ethics.

What has been interesting is the Twitter hatred I've been getting about the article. Some people feel defensive because I didn't acknowledge their work. And in some cases I didn't and I should have. But my point was that we are not reaching engineering students. They are not taking ethics. The very people that we need to warn. And secondarily, it is not trickling up to the policy level. So you might know exactly what's wrong with a particular algorithm but if no one in Washington ever hears it then you have to do better.

**As far as I'm concerned we are in a war. And on the other side are the lobbyists. And we are totally outflanked.**

If you want to build a lobbying firm, who pays for lobbying? It's corporations or foundations. Foundations are more and more being created by rich financiers and technologists. And corporations themselves are not going to be pushing back against power of Facebook and Google because that's how they make money. So, I'm not trying to pick on academia because I think they are flush and aren't doing anything. I'm picking on academia because they are literally the only people that *can* do anything.

Journalists, too. Journalists are essentially the only voice. But they are not technically trained. We are just so overpowered by the lobbyists.

To give you an example: I went to a Congressional subcommittee meeting in 2013 on big data & analytics which I blogged about. The only people that spoke to Congress were lobbyists. And there was never a negative word spoken about big data. There was no dark side. They chose all cherry picked examples about how premature babies were being saved by big data. I said to myself, "I want to be one of the people speaking to Congress". Or at least I want an advocate who sees the social justice angle to be there.

Congress doesn't know about big data and that means we need to educate them. We need to speak a language that they or their staffers can understand. And the only people that are currently doing that are being paid by Big Tech.

**What is next on the horizon for you?**

I started a company called ORCAA to audit algorithms. I have a few clients now which is exciting. I want to get experience making algorithms less racist and less unfair. And then I can build a tool that I can handover to a regulator once they are ready to enforce the laws.

So far I have two small companies as clients that have a reason to want to show to the world that what they are doing is fair. They want to show that their algorithms are fair and transparent. But I'm talking to a bunch of larger potential clients to whom I'm explaining that if they do this wrong then they are open to litigation risk. And since they are large companies that scares them a lot. We don't have a bill of rights yet but we do have litigation risk.

In the meantime, I'm a communicator. I've been going around and speaking and writing op-eds people hate (*she laughs*).

[Data Science](#)[Data Justice](#)[Algorithms](#)[Data Policy](#)[Data Ethics](#)[About](#) [Help](#) [Legal](#)

Get the Medium app

