# Response to Cathy O’Neil’s *Weapons of Math Destruction*

After reading Cathy O’Neil’s [Weapons of Math Destruction](https://www.amazon.com/Weapons-Math-Destruction-Increases-Inequality/dp/0553418815), I put down a few comments. Overall, I was impressed by O’Neil’s insight, honesty, sense of responsibility, and ethical focus. Her writing style is to the point, yet wrapped in artistic expression. To get going, she reveals a fundamental principle and problem with model creation: Models are opinions embedded in mathematics. By making choices over the data we collect, the questions we ask, and what data to include or exclude eventually, we impose our ideologies on the models we build. Often we also teach machines how to discriminate, and from then on they perform this disservice extremely efficiently. We allow our machines to capture poisonous prejudices from the past which embed within them the practice of being unfair.

What was also striking to me was the widespread use of proxies instead of using the down-to-earth fundamental data. When proxies are used the data collection effort is usually much cheaper and simpler. However, models created from proxies are much easier to game or to exploit. It is just simpler to manipulate proxies than the complex reality represented by them. For example, auto insurers often use proxies like high school GPA and credit scores as predictors for driving risk, but omits the all-important drunk driving record. O’Neil also reflects on the use of really bad science when correlation is confused with causation, for example in the case of Frederick Hoffman’s 1896 report which basically declares that race is a strong predictor of life expectancy.

O’Neil defines a “Weapon of Math Destruction” as having all three of the elements: *Opacity*, *Scale*, and *Damage*. I think this is an apt definition and these elements collectively shine light on just about every negative thing that could be said about the application of data science in society. Before reading this book, I was mostly unaware of the practices of gross unfairness and exploitation of the value of personal data going on. The uneasy part to me is how we have to make our case to a machine that has been infused with all kinds of prejudices and inaccuracies. We have no insight into the inner workings of this machine, and it does not help to ask either. There does not seem to be a human available to “rescue” us. If we do find a human, the response is usually that it is company policy and “fair” to abide by the machine’s decision. I started experiencing this sense of unease ever since telephone answering machines started making their appearance. There were some workarounds though, I used to just type a bunch of zeros and eventually I got to speak to a human. This book’s content takes the situation much further.

Some of the WMDs O’Neil mentions are recidivism models like the LSI-R, the risk model attached to mortgage-backed securities during the great recession meltdown, the U.S. News & World Report system of rating colleges, personality tests in hiring departments, scheduling software, e-scores (unlike the FICO scores they resemble), and micro-targeting (for example, in the political landscape). I like her suggestion that we should reframe the question of fairness: instead of squabbling over which metric should be used to determine the fairness of an algorithm, we should instead identify all the stakeholders and then consider and weigh the relative harms associated with each of them. The WMDs she mentions certainly do a lot of harm to many stakeholders. I think it makes a lot of sense to try and minimize the harm as much as possible.

It is a good suggestion that we should regulate the mathematical models and algorithms that continue to have an ever-increasing impact on our everyday lives. Not only should the algorithms be held accountable but also the data scientists that construct them. We do need to measure the impact of algorithms and conduct algorithmic audits. Quality assurance practices need to be applied to data science too. As the author puts it, “put the science into data science.”

O’Neil states that the government has a powerful regulatory role to play. The government can begin by enforcing relevant laws that are already in place (FCRA, ECOA, ADA, HIPAA), and also work to make these laws stronger in the area of the application of algorithms. She even suggests that we could follow the European model which demands that all data collected must be approved by the user in the form of an opt-in. In addition, the reuse of data for other purposes is illegal, in other words, user data may not be sold. Another point made by O’Neil is that models that have significant impact on us, for example credit scores and e-scores, should be open and available to the public with the right to request data modification in case of errors. All of these points make a lot of sense to me. The time is ripe for the government to take action. Exploitation of the value of data obtained from the general public, but in particular from less privileged citizens, has gone on far too long.