7-1-2018

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy

Harris Mateen

Follow this and additional works at: https://scholarship.law.berkeley.edu/bjell

Part of the Law Commons

WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY

Cathy O'Neil. Broadway Books, 2016. 268 Pages.

Harris Mateen[†]

In 2008, Cathy O'Neil, a data scientist at the hedge fund D.E. Shaw, watched as the entire financial sector around her imploded.[1] The financial crisis gave O'Neil, a self-proclaimed math nerd with a Ph.D., a front row seat to the weaponization of mathematics. During the financial crisis, complex mathematical models designed and marketed to be superior to human ability and beyond the confines of human subjectivity turned out to be farces—layers of false assumptions coated in difficult-to-understand layers of mathematics. In order to dupe both the public and themselves, financial analysts at banks used models and algorithms as a justification for their subjective and flawed inputs.

Ten years after the financial crisis, O'Neil sees these nefarious models infiltrating every aspect of modern life. These "Weapons of Math Destruction"—as O'Neil calls them—are (1) "opaque," (2) "beyond dispute or appeal," and (3) disproportionally impact the underprivileged.[2] These Weapons of Math Destruction, or WMDs, replace the role of traditional subjective decision makers and cause those affected by their decision-making to adjust their ways of life to the models. In *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, O'Neil highlights different WMDs that have infiltrated various areas of American life.

† .   J.D. 2020 (U.C. Berkeley).

1.   CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 2 (2016).

2.   *Id.* at 3.

Weapons of math destruction plague a wide array of American life, including the workplace. Before examining their role in employment, it is worth noting the expanse of Big Data's reach. O'Neil's description of their role in education and law enforcement are illustrative. In higher education, for-profit universities such as Corinthian College use online models to send targeted online advertisements that prey on "isolated" and "impatient" individuals with "low self-esteem," and who are "unable to see and plan well for [the] future."[3] This allows predatory diploma mills to target the vulnerable with over-priced programs inferior to those offered at most cheaper community colleges.[4] In law enforcement, predictive policing software found in computer models such as PredPol—while successful in reducing serious crimes—can lead to the over-targeting of vulnerable communities when used to focus on "nuisance crimes" that occur in impoverished areas.[5]

To demonstrate the new role of modeling within the employment sector, O'Neil recounts the story of Kyle Behm. Behm, a young adult, dropped out of Vanderbilt University due to struggles with bipolar disorder.[6] After his condition improved, Behm applied for a minimum wage part-time job at a Kroger grocery store.[7] Behm never received a call back and later learned from a friend that he had been rejected from the job after failing a "Five Factor Model" personality test that appeared remarkably similar to the personality tests he received in the hospital.[8] Kroger's test asked subjective and difficult questions such as: "Which adjective best describes you at work, unique or orderly?"[9] There was no "all of the above" option.[10] The model categorized those who answered "unique" to be "narcissis[tic]" as opposed to those answering "orderly," who were categorized as "conscientious."[11] The opacity and arbitrariness of this and other similar models are a threat to anyone who does not meet the prerequisite attributes set by the model. Every subsequent retail job Behm applied for required him to complete similar tests. Behm was rejected from each of these jobs.[12] It appeared that computer software analyzed Behm's answers and "blackballed" him from the job market.[13]

---

3.   *Id.* at 71.
4.   *Id.*
5.   *Id.* at 86–87.
6.   *Id.* at 105.
7.   *Id.*
8.   *Id.* at 106.
9.   *Id.* at 110.
10. *Id.*
11. *Id.*
12. *Id.* at 106.
13. *Id.*

Applying for jobs has traditionally been a subjective experience. A hiring agent either consciously or subconsciously sizes up a candidate. Do they look right? Talk right? How are they dressed? This subjective nature often allows inappropriate biases to leak into the hiring process. In response to this, legislators implemented and began enforcing employment discrimination laws including the Equal Pay Act of 1963, Title VII of the Civil Rights Act of 1964, and the Americans with Disabilities Act of 1990.

Technologies offered by companies such as Kronos Inc.—which developed the software that ultimately blacklisted Kyle Behm—provide employers with a veneer of fairness.[14] After all, it would seem that the best way to cure human subjectivity is to remove the humans from the decision-making process altogether. As employers seek to be more fair and efficient in their hiring process, more choose to outsource their hiring processes to automated models. According to O'Neil, automation in the hiring process has become a $500 million per year industry.[15] Unfortunately, while this may make the hiring process easier for human resources departments that no longer need to filter through stacks of resumes, it can unfairly push people like Behm out of the workforce. Human biases vary across hirers and are subject to employment discrimination laws. Bad decisions can often be appealed. On the other hand, biases baked into computer systems and used across industries can silently threaten entire populations with unclear repercussions.

Kyle Behm's father—Roland Behm, a prominent Atlanta attorney—took note of this and notified several corporations of his intent to file a class-action lawsuit "alleging [that] the use of the [test] during the job application [is] unlawful."[16] It is unclear whether his impending suit will have any merit. However, in one instance, Rhode Island regulators reprimanded CVS Pharmacy for "illegally screening out applicants with mental illnesses" after conducting a vague personality test.[17]

While what occurred to Behm is particularly egregious, the effects of WMDs in the application process extend beyond individuals with disabilities. Today, according to O'Neil, "72 percent of resumes are never seen by human eyes."[18] This 72 percent instead gets vetted and weeded out by computer models that use their own sets of criteria to score applications for certain jobs.[19] These criteria can include everything from font style and size to keywords such as "summa cum laude" or "sales manager."[20] The

---

14. *Id.* at 110.
15. *Id.* at 108.
16. *Id.* at 106.
17. *Id.*
18. *Id.* at 114.
19. *Id.*
20. *Id.* at 115.

result is that savvy applicants know to litter their resumes with as many keywords as can be picked up by these "automatic readers."[21] Those with knowledge of how the systems operate can better game their resumes for callbacks, but those without this insider tip likely lose their chances of having their resumes read by a human eye. Nevertheless, the livelihood for people in both groups "increasingly [depends] on [their] ability to make [their] case to machines."[22]

For many Americans, the machine dependency does not end after the hiring period. Efficiency-focused work-shift scheduling WMDs increasingly wreak havoc on the lives of low-wage workers. These WMDs analyze upcoming trends—for example, football games in the area, Twitter posts demonstrating increased consumer interest over a weekend, and changes in weather conditions—and use this information to allocate employees' hours as efficiently as possible.[23] The result is that "two-thirds of food service workers and more than half of retail workers find out about scheduling changes with notice of a week or less—often just a day or two."[24] Organizing day-to-day lives becomes difficult for these employees, as they must constantly scramble to rearrange plans after learning their new shift hours. This kind of scheduling could have long-term repercussions on the children of those employed in low-waged jobs. O'Neil quotes the Economic Policy Institute: "Young children and adolescents of parents working unpredictable schedules or outside standard daytime working hours are more likely to have inferior cognition and behavioral outcomes."[25] She believes that the "oversupply of low wage labor" and shrinking of union activity over recent decades have left employees with little bargaining power to push back on the software's subjective decision-making processes.[26]

Finally, even when an employee has made it through the hiring process and locked down their hours, they may now be subject to monitoring and assessment by machines. This monitoring has spread across several industries. In 2008, a now-defunct company called Cataphora created software that attempted to rank employees' "generation of ideas."[27] The software tracked employee email messages looking for ideas that would spread quickly and widely among the organization.[28] After analysis and tracking, the system categorized and ranked employees as "idea generators"

---

21. *Id.* at 114.
22. *Id.*
23. *Id.* at 125.
24. *Id.* at 125–26.
25. *Id.* at 129.
26. *Id.* at 128.
27. *Id.* at 130.
28. *Id.* at 131.

and idea "connectors."[29] This software can be a useful supplement to understanding the roles that employees may play within an organization. However, after the financial crisis, human resources departments needing to lay off staff consulted Cataphora software to determine which employees were most expendable.[30] Cataphora software categorized how employees generate ideas through a single system. O'Neil stresses, however, that Cataphora would be unable to understand factors beyond its system, such as whether employees effectively spread ideas verbally or diffuse office tension through personality traits.[31] Effective employees with these merits that are not regimented into Cataphora's world may have been laid off based on incomplete criteria. As the model cannot track former employees in their new employment, there was no feedback mechanism for the model to learn whether it was right to recommend laying off certain employees.[32] As O'Neil writes, "managers assume[d] that the scores [were] true enough to be useful, and the algorithm [made] tough decisions easy."[33]

Even though Cataphora sold itself to a startup company in 2012, subjective software models that appear objective at first blush still pervade the workplace.[34] Managers increasingly make firing decisions based on often-arbitrary model results. This can lead to a perversion of job responsibilities in order to appease the machine decision-maker. In 2007, education reform advocate Michelle Rhee became the new Chancellor of District of Columbia Public Schools (DCPS).[35] In an attempt to bring accountability into Washington D.C.'s failing public schools, she implemented a program called the DCPS Effectiveness Assessment System for School-Based Personnel (IMPACT), which rated teachers by comparing their students' year-end standardized test scores with the previous years' results.[36] This resulted in teachers cheating in order to maintain their jobs. A *Washington Post* investigation revealed that in some schools, over seventy percent of teachers were suspected of altering their students' standardized tests.[37] As more models become firing tools, more employees will start to do their jobs in ways to meet the model's qualifications rather than in ways that may be more effective, but are difficult for the model to capture.

The use of algorithms within the employment sphere is often paired with good intentions, such as trying to make workplaces fairer. One startup,

---

29.   *Id.*
30.   *Id.* at 132.
31.   *Id.* at 133.
32.   *Id.*
33.   *Id.*
34.   *Id.* at 133.
35.   *Id.* at 3–4.
36.   *Id.*
37.   *Id.* at 9.

Gild, claims to find candidates for technology companies from more socioeconomically diverse backgrounds.[38] The push to digitize hiring to foster increased diversity in workspaces is rooted in empirical studies indicating biases among human hirers.[39] One study on law firm hiring discovered that firms tend to view resumes signaling upper-class backgrounds to be better "fit[s]" than resumes signaling poorer backgrounds.[40] Startups focused on hiring often claim their software can cut back on this human bias. Nevertheless, blind faith in these models can lead to overreliance on systems that may have bias unintentionally baked into them. At worst, a well-intentioned startup may accidentally create a Weapon of Math Destruction.

One way to prevent this from occurring, O'Neil states, is to implement a broad regulatory system, both internally within organizations and externally at government enforcement agencies. The regulatory system should hunt for "hidden costs" within widely used algorithms.[41] One component of this regulatory regime involves "algorithmic audits."[42] Data scientists would study the outputs of major algorithms to determine the inputs incorporated by the model.[43] According to O'Neil, some egregious models that cannot ensure compliance may need to be scrapped.[44] She also suggests that in situations where efficiency-focused models—such as the work-shift schedulers—overburden the victims of its actions, we may need to "dumb down our algorithms."[45]

Cathy O'Neil herself has founded O'Neil Risk Consulting & Algorithmic Auditing (ORCAA), an algorithmic risk consulting and auditing firm, which bills itself as helping companies avoid litigation risk resulting from unfair algorithms.[46] As almost all companies begin to incorporate modeling technology into their operations, algorithmic auditing firms—which barely exist today—have the potential to be as widespread as financial auditing firms such as Ernst & Young or Deloitte.

However, while necessary, algorithmic auditing probably cannot scale without proper incentive structures for software companies to participate. Currently, major tech players such as Google and Facebook refuse to open

---

38.   Claire Cain Miller, *Can an Algorithm Hire Better Than a Human?*, N.Y. TIMES (June 25, 2015), https://www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html.

39.   *See* Lauren A. Rivera & András Tilcsik, *Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market*, 81 AM. SOC. REV. 1097 (2016).

40.   *Id.*

41.   O'NEIL, *supra* note 1, at 207.

42.   *Id.* at 208.

43.   *Id.*

44.   *Id.*

45.   *Id.* at 210.

46.   O'NEIL RISK CONSULTING & ALGORITHMIC AUDITING, http://www.oneilrisk.com (last visited Nov. 9, 2018).

their models to outsiders.[47] This makes sense since models such as Google's search algorithm are amongst the most valuable pieces of proprietary technology in the world. Anything that could compromise these models could cost companies billions of dollars. Nevertheless, if these models disproportionately harm people with protected traits, some sort of mechanism should exist to understand and correct them. It is unclear whether any auditing mechanism can exist that both maintains the shroud of secrecy around a company's technology while also ensuring accountability.

Ensuring the use of algorithmic auditing would likely require amending current regulation to account for the effects of modern technology in the American workplace. For example, O'Neil suggests updating the American with Disabilities Act (ADA) to "take into account Big Data personality tests."[48] It seems obvious to adapt existing law to the effects of rapidly changing technology. Nevertheless, resolution of this depends on Congress's willingness to amend laws such as the ADA. This seems to be a distant solution so long as the U.S. Congress remains allergic to regulation.

Another more immediate approach may be to bring lawsuits similar to the one Kyle Behm's father intends to file against companies inappropriately using WMDs. However, it is unclear how successful such lawsuits will be without first amending the current law. In a 2016 Wisconsin criminal case, defendant Eric Loomis claimed his due process rights were violated when a judge used Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), a risk assessment tool developed by a private entity, to assist in sentencing Loomis.[49] Part of COMPAS's model incorporated Loomis's gender. Loomis argued that gender was a "criminogenic factor," whereas the prosecution claimed that gender was not a "criminogenic factor" and was instead used for statistical "norming."[50] Neither group could be certain "due to the proprietary nature of COMPAS."[51] The Wisconsin Supreme Court ultimately found no violation of Loomis's due process rights in the lower court's use of COMPAS for sentencing. Although not a civil case, *Loomis* demonstrates a court's willingness to use a tool's recommended outcome without a complete understanding of how the tool arrived at the outcome. If applied in civil litigation, it does not present a hopeful sign to curbing Weapons of Math Destruction.

O'Neil's *Weapons of Math Destruction* paints an ominous vision of the present and an even eerier vision of the future. Workers, no longer at the

---

47.   O'NEIL, *supra* note 1, at 211–12.

48.   *Id.* at 213.

49.   *See* State v. Loomis, 371 Wis.2d 235, 244 (Wis. 2016).

50.   *Id.* at 267.

51.   *Id.*

whims of subjective employers, are now subject to the harsher whims of opaque models paraded as objective decision-makers. As the nature of our economy changes and Weapons of Math Destruction continue to proliferate throughout the employment landscape, I hope that we as a nation will be able to confront this phenomenon and find a policy solution.

Harris Mateen, J.D. 2020 (U.C. Berkeley)