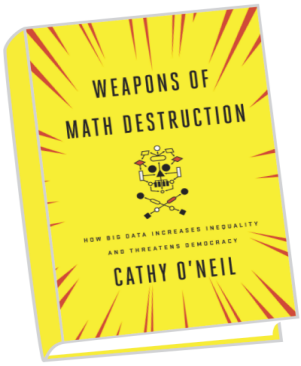
BOOK REVIEW



Weapons of Math Destruction

Reviewed by Mary Poovey

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy
by Cathy O'Neil
Crown, New York, 2016
259 pages

Cathy O'Neil's *Weapons of Math Destruction* is a timely reminder of the power and perils of predictive algorithms and model-driven decision processes. The book deals in some depth with eight case studies of the abuses she associates with WMDs: "weapons of math destruction." The cases include the havoc wrought by value-added models used to evaluate teacher performance and by the college ranking system introduced by *U.S. News and World Report*; the collateral damage of online advertising and models devised to track and monetize "eyeballs"; the abuses associated with the recidivism models used in judicial decisions; the inequities perpetrated by the use of personality tests in hiring decisions; the burdens placed on low-wage workers by algorithm-driven attempts to maximize labor efficiency; the injustices written into models that evaluate creditworthiness; the inequities produced by insurance companies' risk models; and the potential assault on the democratic process by the use of big data in political campaigns. As this summary suggests, O'Neil had plenty of examples to choose from when she wrote the book, but since the publication of *Weapons of Math Destruction*, two more problems associated with model-driven decision procedures have surfaced, making O'Neil's work even more essential reading. The first—the role played by fake news, much of it circulated on Facebook, in the 2016 election—has led to congressional investigations. The second—the failure of algorithm-governed oversight to recognize and delete gruesome posts on the Facebook Live streaming service—has caused CEO Mark Zuckerberg to announce the addition of 3,000 human screeners to the Facebook staff. While O'Neil's book may seem too polemical to some readers and too cautious to others, it speaks forcefully to the cultural moment we share.

O'Neil weaves the story of her own credentials and work experience into her analysis, because, as she explains, her training as a mathematician and her experience in finance shaped the way she now understands the world. O'Neil earned a PhD in mathematics from Harvard; taught at Barnard College, where her research area was algebraic number theory; and worked for the hedge fund D. E. Shaw, which uses mathematical analysis to guide investment decisions. When the financial crisis of 2008 revealed that even the most sophisticated models were incapable of anticipating risks associated with "black swans"—events whose rarity make them nearly impossible to predict—O'Neil left the world of corporate finance to join the RiskMetrics Group, where she helped market risk models to financial institutions eager to rehabilitate their image. Ultimately, she became disillusioned with the financial industry's refusal to take seriously the limitations of risk management models and left RiskMetrics. She rebranded herself a "data scientist" and took a job at Intent Media, where she helped design algorithms that would make big data useful for all kinds of applications. All the while, as O'Neil describes it, she "worried about the separation between technical models and real people, and about the moral repercussions of that separation" (page 48). O'Neil eventually left Intent Media to devote her energies to in-

The book speaks forcefully to the cultural moment we share.

Mary Poovey is Samuel Rudin University Professor in the Humanities Emeritus and professor emeritus of English at New York University. Her e-mail address is mary.poovey@nyu.edu .

DOI: http://dx.doi.org/10.1090/noti1561

Evolving from academic mathematician to quant to blogger, Cathy O'Neil has accumulated an unusual set of experiences and expertise. Her book reviewed here, *Weapons of Math Destruction*, was long-listed for the National Book Award. Her earlier book *Doing Data Science*, with co-author Rachel Schutt, was reviewed by Brian Hayes in the October 2014 issue of the *Notices*.[1] O'Neil blogs at https://mathbabe.org.

vestigating and blogging about what she now considers the most serious side effect of our reliance on algorithm-driven evaluation: economic inequality. In what she calls a "growing dystopia," "the algorithms…make sure that those deemed losers…remain that way. A lucky minority… gain ever more control over the data economy, raking in outrageous fortunes and convincing themselves all the while that they deserv[e] it" (page 48).

The sequence of O'Neil's case studies roughly follows the life stages of many middle- and working-class individuals. The impact of the value-added evaluation model on a fifth-grade teacher provides an introduction to our prevailing obsession with ranking procedures, whose unanticipated consequences include the high tuition costs of American colleges and universities, which stem, in part, from expensive campaigns designed to raise a college's standing in the *U. S. News and World Report* annual survey. Her treatment of online advertising begins with an analysis of for-profit colleges like the (now defunct) Corinthian College, then moves to the practice of "lead generation," which amplifies the predatory nature of internet advertising targeted to aspiring, but poorly funded, students. Encounters with the justice system, the job market for low-wage work, shift-work at fast food restaurants and big box stores, payday lenders, the insurance market, and campaigns designed to lure people to the ballot box (or frighten them away) constitute all-too-familiar, nearly inescapable episodes for millions of Americans. While O'Neil's argument is explicitly political—in keeping with the Occupy Wall Street movement, in which she participated—it is important to point out that her case studies illuminate problems everyone faces in a society increasingly dominated by algorithm-driven evaluation. You don't have to be poor to feel you are being stalked by predatory marketers on the World Wide Web. And, as the 2016 election made clear, you don't have to be uneducated to be inundated with fake news on the social media platforms.

O'Neil is careful to distinguish between the socially beneficial potential of judiciously constructed and carefully monitored mathematical models and the damage inflicted by models that have become mathematical Weapons of Mass Destruction. She is not, in other words, hostile to mathematics, modeling, or big data. She considers mathematical models "the engines of the digital economy" (page 211). At the end of the book she introduces several examples of models created for the common good: a model that seeks to understand the conditions that push homeless families back into shelters and a model that helps identify components in products that have been made by enslaved labor, which is presumably designed to encourage manufacturers to reject those components or improve working conditions throughout the supply chain. By and large, however, O'Neil focuses on what makes a model a Weapon of Math Destruction and the destruction these models unleash. According to O'Neil, a WMD is characterized by three features: opacity, scale, and the damage the model causes. O'Neil acknowledges that every mathematical model articulates the opinions of its creators. As she puts it, "Models are opinions embedded in mathematics" (page 21). But, as a special kind of model, the WMD hides its foundational assumptions in an impenetrable black box. Such models obscure the source and kind of their input information and model parameters, they rely on proxy data instead of directly observable inputs, and they create invisible feedback loops that make their effects nearly inescapable. These models scale easily and are often implemented in populations or situations for which they were not designed. "The developing WMDs in human resources, health, and banking, just to name a few, are quickly establishing broad norms that exert upon us something very close to the power of law," she explains. "If a bank's model of a high-risk borrower, for example, is applied to you, the world will treat you as just that—a deadbeat—even if you're horribly misunderstood. And when that model scales, as the credit model has, it affects your while life—whether you can get an apartment or a job or a car to get from one to the other" (page 30). As this passage suggests, the damage caused by these models can be enormous—both for the individual and for society as a whole. Just consider the percentage of the US prison population composed of African American, often poor or working-class men, many of whom languish behind bars because recidivism models predicted they would commit additional crimes after being arrested for what was often a minor initial offense.

大规模杀伤性武器有三个特征：不透明度、规模和模型造成的损害

模型是嵌入数学中的观点

"Models are opinions embedded in mathematics."

As much as I admire *Weapons of Math Destruction*, I can imagine two kinds of objections to this book. The first

[1] www.ams.org/notices/201409/rnoti-p1068.pdf

would argue that mathematics is an objective science and that to blame mathematical models for contributing to social inequality is simply wrong-headed. ~~Corporations intent on maximizing profits might misuse mathematical models, this argument would go, but neither the algorithm nor big data is inherently predisposed to privilege one social group over another.~~ To a certain extent, I suspect that O'Neil would agree with this position. After all, she is clearly committed to mathematics as an instrument of inquiry and a tool that can be used for social improvement, and she explicitly admires the application of models that do not morph into WMDs—models that are continuously checked against new inputs and updated accordingly, such as models used to promote social improvement or even the *Moneyball* models used to select players in professional sports. To a certain extent, O'Neil anticipates and answers the objections that mathematicians might mount, but she does not back down from cautionary warnings about overconfidence in mathematical objectivity. She insists that mathematical models are insufficiently subtle to capture concepts, attitudes, and ideas; she warns that every model omits some components of a situation in order to highlight others; and she reminds us that it can be too easy to move models developed for one situation to another where they do not apply. O'Neil also notes that the success (some) models have had in predicting (some kinds of) events has paradoxically made modeling as a whole both more ubiquitous and more dangerous, as those who design and use models are easily lulled into confidence that their models are more neutral, robust, and scalable than events can prove them to be.

Discrimination and discernment do not yield objective results and resist machine learning.

Even if O'Neil's book anticipates this first kind of objection, she does not fully appreciate a second line of argument, which—to me at least—is even more important. This would come not from mathematicians but from humanists (like myself) or from critics of neoliberalism (such as the economist Philip Mirowski). The series of challenges this group might mount would argue that O'Neil's criticisms of WMDs do not go far enough. While she exposes the collateral damage inflicted by models that are too opaque, large, and malicious, she does not fully convey the price our society will pay if we continue to equate evaluation exclusively with measurement. O'Neil valiantly urges readers to examine our social norms: "We must reevaluate our metric of success" (page 206), she argues. But she does not dwell on—or sufficiently credit—the process of judgment by which, throughout history, human beings have collectively decided what constitutes success in the first place. Teachers in the humanities routinely try to cultivate in students the difficult skills of discrimination and discernment, comparison and appreciation that algorithms are designed to replace. All of us know these skills do not yield objective results; they resist machine learning, require extensive time and labor, and generate norms that require continuous, often contentious, conversation and debate. Yet we continue to value these skills because we believe the outcomes they generate are ultimately more fair because they are more fully considered. Some of us worry that, as evaluation is increasingly equated exclusively with numerical measurement, we will allow these more time-consuming skills to atrophy. The loss—or devaluation—of these skills would not only make the oversight of mathematical modeling that O'Neil calls for increasingly difficult; it would also present ever-larger challenges to those who want to formulate the kind of critique O'Neil offers here.

Among the critiques that will be swept aside if we continue to prefer numerical measurement to discrimination is the charge that the avalanche of data that now so radically shapes our lives is driven not simply by inevitable advances in computing and software but by a neoliberal corporate campaign designed to dominate Western society. Philip Mirowski makes this argument forcefully in *Never Let a Serious Crisis Go to Waste*, but it has been made by others, including members of the Occupy movement. Whether or not one believes that privileging efficiency over fairness articulates a coherent corporate ideology associated with the One Percent, it is easy to see that pursuing this or any other critical line of argument will become increasingly difficult as the value we collectively ascribe to critical thinking diminishes. O'Neil recognizes that the choices embedded in models are "fundamentally moral" (page 218), and she brilliantly exposes some of the consequences of the decisions model builders make. But O'Neil does not register the full extent of what we stand to lose as we transfer the very processes of judgment from humans to algorithms.

Photo Credit

Photo of Cathy O'Neil courtesy of Cathy O'Neil.

ABOUT THE REVIEWER

Mary Poovey is the author of many books, including the forthcoming *Finance in America: An Unfinished Story*, co-authored with Kevin R. Brine. Her article "Can Numbers Ensure Honesty? Unrealistic Expectations and the U.S. Accounting Scandal," based on a public lecture she presented at the International Congress of Mathematicians in Beijing in 2002, appeared in the January 2003 issue of the *Notices*.[2]

[2] www.ams.org/notices/200301/fea-poovey.pdf