

Weapons of Math Destruction

Author: O’Neil, Cathy

“Weapons of Math Destruction,” by statistician Cathy O’Neal, is about how improperly using mass data can punish the marginalized (e.g. In terms of finances, education, criminal justice, etc.).

An algorithm processes a slew of statistics and comes up with a probability that certain person might be a hire, a risky borrower, a terrorist, or a miserable teacher. That probability is distilled into a score. The human victims of WMDs are held to a far higher standard of evidence than the algorithms themselves. (10)

Profits end up serving as a stand-in, or proxy, for truth. (12)

“Moneyball” is now shorthand for any statistical approach in domains long ruled by the gut. (16) It is a model that is an abstract representation of some process. (18) Models are, by their very nature, simplifications. To create a model, then, we make choices about what’s important enough to include, simplifying the world into a toy version that can be easily understood and from which we can infer important facts and actions. (20) A model’s blind spots reflect the judgments and priorities of its creators. Models, despite their reputation for impartiality, reflect goals and ideology. Models are opinions embedded in mathematics. Whether or not a model works is also a matter of opinion. A key component of every model, whether formal or informal, is its definition of success. (21) A model built for today will work a bit worse tomorrow. (22)

Three components of WMD: Opacity; Scale; Damage.

Mathematical models, by their nature, are based on the past, and on the assumption that patterns will repeat. (38) Developers of models often wield formulas to impress rather than clarify. (44)

WMDs create a vicious feedback loop where the ranking becomes destiny. (53)

When you create a model from proxies, it is far simpler for people to game it. This is because proxies are easier to manipulate than the complicated reality they represent. (55)

Ask: What is the objective of the modeler? A spectacular failure comes from what they chose not to count. (59) It is much safer to start with the venerable champions on top. (60)

WMDs result in systems that favor the privileged. (85)

Anywhere you find the combination of great need and ignorance, you’ll likely see predatory ads. (70) The key for an optimization program, naturally, is to pick an objective. (73)

The Constitution sacrifices enormous efficiencies for the promise of fairness. The Constitution’s implicit judgement is that freeing someone who may well have committed a crime, for lack of evidence, poses less of a danger to our society than jailing or executing an innocent person. WMDs, by contrast, tend to favor efficiency. (85) A crucial part of justice is equality. Justice cannot just be something that one part of society inflicts upon the other. (96)

From a mathematical point of view, trust is hard to quantify. That's a challenge for people building models. (103)

Question how automatic systems judge us. (107) An example: In customer service, talking was discouraged because workers were supposed to spend as many of their minutes as possible on the phone, solving customers' problems. Coffee breaks were scheduled one by one. The researchers found, to their surprise, that the fastest and most efficient call center team was also the most social. (132)

Simpson's Paradox: When a whole body of data displays one trend, yet when broken into subgroups, the opposite trend comes into view for each of those subgroups. (136)

The impact of value-added modeling is pervasive and nefarious. (139)

FICO scores have a clear feedback loop and are relatively transparent. E-scores, conversely, codify past injustices into their models by using zip codes to classify. The modelers for e-scores have to make do with trying to answer the question, "How have people like you behaved in the past?" when ideally, they would ask, "How have YOU behaved in the past?" The difference between these two questions is vast. (146)

Framing debt as a moral issue is a mistake. (149) The unregulated side of the data economy is even more hazardous. (151) We can't count on automatic systems to address the issue. (155)

Insurance grew out of actuarial science. Mathematicians didn't pretend to foresee the fate of each individual. But they could predict the prevalence of accidents, fires, and deaths within large groups of people. (163) Surveillance will change the very nature of insurance. (170)

Much of the proxy data collected is not protected by law, so it would theoretically be perfectly legal. (175)

Example of proxy data is the BMI score, which is a person's weight in kilograms divided by the square of their height in centimeters. It's a crude numerical proxy for physical fitness. It's more likely to conclude that women are overweight. (186) A key question about proxy data is whether this help is an offer or a command. (177)

Emotional states can be transferred to others.... leading people to experience the same emotions without their awareness. Facebook's algorithms can affect how millions of people feel, and those people won't know that it's happening. People widely trust search engines. Some 73 percent of Americans, according to a Pew Research report, believe that search results are both accurate and impartial. (184)

Microtargeting. As it happens, it will become harder to access the political messages our neighbors are seeing – and as a result, to understand why they believe what they do, often passionately. The result of these subterranean campaigns is a dangerous imbalance. The political marketers maintain deep dossiers on us, feed us a trickle of information, and measure how we respond to it. But we're kept in the dark about what our neighbors are being fed. The asymmetry of information prevents the various parties from joining forces – which is precisely the point of the democratic government. (185)

Our national motto, E Pluribus Unum, means “Out of Many, One.” But WMDs reverse the equation. Working in darkness, they carve one into many, while hiding us from the harm they inflict upon our neighbors near and far. (200)

WMD victims are nearly voiceless. WMDs generating fabulous profit margins are not likely to remain cloistered for long in the lower ranks. That’s not the way markets work. They’ll evolve and spread, looking for new opportunities. And they’ll continue to multiply, sowing injustice, until we take steps to stop them. (203)

Big Data processes codify the past. They do not invent the future. Doing that requires moral imagination, and that’s something only humans can provide. In a sense, our society is struggling with a new industrial revolution. And we can draw some lessons from the last one. (204)

To eliminate WMDs, we must advance beyond establishing best practices in our data guild. Our laws need to change, too. And to make that happen we must reevaluate our metric of success. (206) A regulatory system for WMDs would have to measure such hidden costs, while also incorporating a host of non-numerical values. This is already the case for other types of regulations. Numbers can never express their value. And the same is often true of fairness and the common good in mathematical models. They’re concepts that reside only in the human mind, and they resist quantification. (207)

There’s no fixing a backward model like the value-added model. The only solution is to ditch the unfair system; incorporate positive feedback loops so that the model can be improved. (209)

Movements toward auditing algorithms. Moving toward the European model, which stipulates that any data collected must be approved by the user, as an opt-in. Models that have a significant impact on our lives, including credit scores and e-scores, should be open and available to the public. (214)

Polling itself can actually affect the outcome of an election. (221)

Ask: What do we mean by “bias”? (222)

In the end, how an algorithm will be used should affect how it is constructed and optimized with a system of accountability. Start putting the “science” into data science. Equip ourselves with sufficient skepticism. It all starts with ground truth. (227)