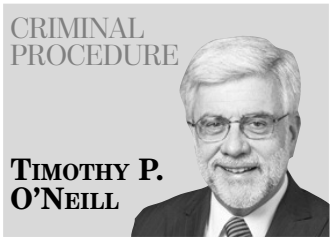
Chicago Daily Law Bulletin®

Volume 162, No. 235 Serving Chicago's legal community for 161 years

Algorithms aim to downplay the human element, but statistics can blindfold justice



Timothy P. O'Neill is the Edward T. and Noble W. Lee Chair in Constitutional Law for 2014-15 at The John Marshall Law School in Chicago. Readers are invited to visit his Web log and archives at jmls.edu/oneill.

One thing we can all agree on is that computers and sophisticated algorithms have enabled us to make much more rational decisions on everything from the cost of insurance to eligibility for a mortgage. An algorithm eliminates human bias by guaranteeing that everyone is treated equally.

Right?

Wrong, says a book written by data scientist Cathy O'Neil and nominated for a 2016 National Book Award. It argues the mathematical models used to create these programs not only reflect the unfairness already found in society, but their use also then creates a "feedback loop" that actually worsens the problem. The book is called "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy" (Crown, 2016).

Consider, for example, an algorithm used to screen potential hires at a company. One factor that could be used is an applicant's credit score. It is a reasonable factor to consider; after all, someone who pays her bills on time is probably more likely to be punctual and reliable.

The problem, of course, is that there are lots of good workers who may experience an economic crisis that causes them to experience a decline in their credit scores.

But, since the algorithm equates bad credit with bad job performance, the good worker who has experienced an economic crisis will be less likely to get a job. Thus, the economically strapped good worker being denied a job will be more likely to fall into poverty. Poverty means she will have trouble paying new bills, and her credit score will continue to fall. In this way, the algorithm has created a feedback loop that perpetuates a divide in society.

For an example of this effect in criminal law, O'Neil discusses an assessment called the Level of Service Inventory Revised, commonly referred to as the LSI-R. It is used in many places around the country, including Illinois. It is a lengthy questionnaire that prisoners are asked to complete. According to the assessment publisher, the results are then used to predict "parole outcome, success in correctional halfway houses, institutional misconducts, and recidivism."

O'Neil notes that it includes a number of relevant inquiries. For example, asking "How many prior convictions have you had?" is clearly germane to the risk of recidivism. And "What part did drugs and alcohol play in your offense?" is information vitally needed to make treatment decisions.

But O'Neil notes the problematic nature of a question such as "When was the first time you were ever involved with the police?"

A prisoner from the suburbs may very well state that it is the crime for which he is presently incarcerated. Yet a young black male from an urban neighborhood is statistically more likely to have been stopped by the police on numerous occasions. (A 2013 study by the American Civil Liberties Union showed that in New York, black and Latino males between 14 and 24 accounted for 40.6 percent of police stops while comprising only 4.7 percent of the city population; more than 90 percent of those stopped were innocent of any wrongdoing.)

So if "early involvement with the police" is a thumb on the algorithm's scale for recidivism, racial minorities from urban areas always look like far greater risks.

O'Neil notes similar problems with a question concerning whether the prisoner has friends or relatives with criminal records.

Based on these responses, the LSI-R score categorizes a prisoner as a high, medium or low risk for recidivism. Again, the problem is that the score creates its own feedback loop. A prisoner labeled high-risk is likely to come from a neighborhood where he, his family, and his friends have all had involvement with the police. This will result in the greater likelihood that he will remain in prison for a longer period of time surrounded by other high-risk prisoners.

When he is eventually released back to his neighborhood, his criminal record will make it much harder for him to find a job. Thus he may decide to commit another crime. And if he does, in the words of O'Neil, "the [LSI-R's] recidivism model can claim another success." The model itself "contributes to a toxic cycle and helps to sustain it." This is why O'Neill refers to these tools as "Weapons of Math Destruction."

So if "early involvement with the police" is a thumb on the algorithm's scale for recidivism, racial minorities from urban areas always look like far greater risks.

O'Neil notes that although human decision-making is flawed, it does have one great virtue: "It can evolve." Humans can learn and adapt and change. But automated systems and algorithms "stay stuck in time until engineers dive in to change them."

An example of this in civil law appeared in an Oct. 25 report in The Washington Post. A jury in Brooklyn found a landlord liable for lead poisoning that resulted in a 4-year-old Hispanic boy becoming disabled. The next issue was to determine damages by using economic models to estimate what he would have earned during his lifetime had he not been disabled. The boy's lawyer argued that $3.4 million was proper since his parents were both highly educated and it could be assumed he would have a prosperous future.

Yet opposing counsel presented statistics indicating the percentage of Hispanics earning a master's degree is only around 7 percent. He therefore argued that the recovery should be only $1.5 million.

The trial judge refused to allow the argument, holding that race and ethnicity should not be used in estimating an individual's future prospects because this "reinforces the rigid racial and ethnic barriers that our society strives to abolish."

Equally problematic is that this use of statistics does not take into account how current gaps will change in the future. It locks in current unfairness. For example, an estimate in 1970 of the future earnings of a 20-year-old female would have underestimated the actual amount by about 28 percent.

O'Neil realizes that although algorithms are here to stay, we have to confront their shortcomings: "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. We [must create] Big Data models that follow our ethical lead. Sometimes that will mean putting fairness ahead of profit."

The challenge is to ensure that data never excludes human values.