Weapons of Math Destruction

INTRODUCTION

BRIEF BIOGRAPHY OF CATHY O'NEIL

Cathy O'Neil was interested in math from a young age. After attending UC Berkley as an undergraduate and earning a PhD in mathematics (with a focus on algebraic number theory) from Harvard University in 1999, she held a postdoc appointment at the MIT math department and a professorship at Barnard College. After leaving Barnard to work as a quant (or quantitative analyst) at D.E. Shaw, a major hedge fund, O'Neil found herself at the center of the 2007-2008 global financial crisis. In 2011, O'Neil left Shaw to work as a data scientist at an e-commerce startup, but she found herself increasingly disillusioned by how faulty and dangerous algorithms had become central to almost every sector of the economy. O'Neil joined the Occupy Wall Street movement and started a blog, *mathbabe*, where she focused on "exploring and venting about quantitative issues." O'Neil is the author of *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* and the founder of ORCAA (O'Neil Risk Consulting & Algorithmic Auditing), a consultancy focused on helping companies and organizations responsibly manage and audit algorithms.

HISTORICAL CONTEXT

Throughout *Weapons of Math Destruction*, author Cathy O'Neil traces the roots of the modern-day data economy. She describes the beginnings of the insurance industry in the 1600s and the early days of data's role in the American justice system in the mid-20th century. She even delves into the obscure but meaningful connection between Big Data and the racist and long-debunked study of phrenology (how bumps and ridges on the human skull were believed to dictate certain traits and characteristics) in the 18th-20th centuries. O'Neil also discusses more modern-day socioeconomic and political happenings, such as the 2007-2008 financial crisis which began on Wall Street but had ripple effects throughout the global economy. Even more recently, she examines the 2016 U.S. presidential election, in which bad polling data played a critical role.

RELATED LITERARY WORKS

Weapons of Math Destruction is one of many books that explore how data and algorithms are increasingly influential in almost every aspect of contemporary life. Michael P. Lynch's *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data* explores how internet algorithms have helped—and hindered—the human learning process. *Weapons of Math Destruction* also touches on some major social and political milestones of the early 21st century, including the 2007 Wall Street crash and the 2016 U.S. presidential election. *Too Big to Fail* by Andrew Ross Sorkin examines the financial crash from the inside out, while Hillary Clinton's *What Happened* examines how polling data and resulting campaign strategy gaffes derailed her run for the presidency. *Weapons of Math Destruction* also examines social media platforms' role in entrenching data and algorithms in every aspect of life. In a similar vein, *The Accidental Billionaires* by Ben Mezrich explores the founding of Facebook, exploring how the company's roots have shaped its current power in the tech world.

KEY FACTS

- **Full Title:** Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy
- **When Written:** Mid-2010s
- **Where Written:** New York City
- **When Published:** 2016
- **Literary Period:** Contemporary
- **Genre:** Nonfiction
- **Climax:** O'Neil describes how flawed and incomplete polling algorithms failed to predict the outcome of the 2016 U.S. presidential election.
- **Antagonist:** "Weapons of math destruction" (WMDs)
- **Point of View:** First Person and Third Person

EXTRA CREDIT

Big Issues. On her popular blog *mathbabe* and in articles for *Bloomberg*, Cathy O'Neil continues to explain how dangerous algorithms are impacting daily life in the modern world. She's written articles on how "TikTok's Algorithm Can't Be Trusted" and how limited access to the COVID-19 vaccine threatens to worsen global inequality, as well as blog posts about the economy and the threat of hyperinflation.

PLOT SUMMARY

In *Weapons of Math Destruction*, mathematician and writer Cathy O'Neil explores the world of Big Data—and its insidious, fast-growing control over almost every aspect of modern life. In order to understand why faulty, invasive new algorithms are so widespread and powerful, O'Neil explores the history of how models have been used since the 17th century to determine things like who can buy insurance, how students learn, how politicians run their campaigns, and what kinds of sentences criminals receive.

O'Neil is a mathematician who has put her skills to use as a professor of mathematics, as a quant for the hedge fund D.E. Shaw, and as a data analyst for numerous start-ups. She became disillusioned with the so-called "Big Data economy" around the time of the 2007-2008 financial crisis. The crisis originated because lenders were using subprime mortgages to create mortgage-backed securities—in simple terms, an entire economy was being built on nothing. These people had misused math—a sacred tool to a passionate mathematician like O'Neil—and destroyed lives in the process. As O'Neil began thinking about data's stranglehold on modern life, she started to investigate the role of mathematical models in some of humanity's most important institutions: schools, insurance companies, the justice system, and more.

By examining how inaccurate teaching assessments and biased recidivism models sacrifice fairness and justice in the name of efficiency, O'Neil suggests that harmful computer algorithms that she terms "**weapons of math destruction**," (WMDs) have taken over what used to be analog, human-driven processes. Instead of a banker meeting with a pair of newlyweds before determining whether their bank should offer them alone, predatory e-scoring models and other WMDs now determine who's fit to receive what kind of loans or credit. And instead of a judge sentencing a person based on the severity of their crime, criminals are now subject to models that size up their family members' and acquaintances' criminal histories, deciding whether they're likely to offend again based on their home environment.

Weapons of math destruction must meet a few main criteria: they must be "opaque," widespread, and damaging. In other words, their methods of gathering data (or the purposes for which they use that data) must be hard to decipher, their influence must be vast, and they must create hardship or deepen inequality in society. To illustrate how destructive these algorithms are, O'Neil explains how WMDs have infiltrated the college admissions process in the U.S. (and college rankings to boot); how they define job hunts and work schedules; how they determine who is eligible for credit cards, loans, and insurance policies; and how they interfere with political campaigns and elections.

After exploring how faulty or biased algorithms threaten much of U.S. society—from policing to school systems to coffee shops—O'Neil concludes that it will be difficult to dismantle WMDs because of how interconnected society is. The same data that encourages for-profit colleges to send out predatory ads and law enforcement to use predictive policing software categorizes people based on the "risk" they represent to society.

There's still time for corporations to right wrongs in their algorithms, remove bias from their models, and restore humanity's hand in making big decisions about the fates of students, workers, and consumers. Machines, unlike humans, have no concept of fairness. WMDs could be used to help people—they could predict spots where child abuse is more likely to occur or stop corporations from using slave labor in their product manufacturing. But instead, they're just making modern life more unequal (and more automated). O'Neil concludes that while there may never be a single definition of what makes an algorithm fair, it's up to corporations and lawmakers to set standards for how to hold algorithms accountable and improve how they work.

CHARACTERS

MAJOR CHARACTERS

Cathy O'Neil – Cathy O'Neil is a mathematician, data scientist, writer, and the author of *Weapons of Math Destruction*. A self-described "math nerd" since childhood, O'Neil took her love of math into the business world when she joined a prominent hedge fund, D.E. Shaw, as a quantitative analyst (or "quant") in the months before the 2007-2008 financial crisis began. Disillusioned by how the players in the crisis had misused math, O'Neil started to think harder about the role that data was playing in everyday life. She was startled by what she found as she started investigating the "Big Data" economy, or the use of personal information to calculate human potential in a variety of arenas. By identifying the term "**weapons of math destruction**" (WMDs) to describe faulty or dangerous algorithms. WMDs, she asserts throughout the book, must be opaque, widespread, and damaging. O'Neil takes a direct and careful approach to describing how WMDs influence the way people around the world live today: WMDs can often define if a person gets into college (and where), whether they're able to secure a job or land an insurance policy, how their workload will be scheduled, and which advertisements we see on the internet. Throughout the book, O'Neil seeks to blow the whistle on how and why companies employ WMDs, the damage these algorithms can do to global society, and what policy changes and review processes are needed to combat them.

Barack Obama – Barack Obama is an American politician who served as the 44th president of the United States from 2009–2017. Barack Obama's political legacy, O'Neil shows, is deeply entwined with the era of Big Data. For instance, his campaign team used analytics experts to create micro-targeting campaigns in the run-up to his 2012 reelection. And, in his second term, Obama passed legislation aimed at creating a new system of college rankings. In these ways, his presidency overlapped with the rise of computer algorithms dictating how people navigate the world.

Mitt Romney – Mitt Romney is an American politician who was the Republican candidate for president during the 2012 U.S.

presidential election. Romney made a serious error at a fundraising dinner in Florida when he invoked right-wing talking points that were covertly filmed and broadcast on social media by event caterers. O'Neil asserts that Romney was speaking based on the "data" he had gathered about the audience he was speaking to at the event, which was largely attended by conservatives—but he'd failed to account for the outliers (the caterers) who would hear his speech as well.

Hillary Clinton – Hillary Clinton is an American politician who was the Democratic candidate for president during the 2016 U.S. presidential election. Clinton's 2016 campaign relied heavily on the use of political polling data to determine strategy—but as a result of inaccurate poll results, Clinton overlooked campaigning in areas of the country where the data seemed to show that she was doing well. Even though nationwide public opinion seemed to indicate that Clinton would win the race against Donald Trump, Trump won the election, which shocked many people. But O'Neil asserts that a rise in populism, media skepticism, and reluctance to contribute data to polls created the illusion that Clinton was the preferred candidate.

Sarah Wysocki – Sarah Wysocki is a Washington, D.C.-area teacher whose career was jeopardized in 2007 by an assessment tool called IMPACT, a data-backed approach to weeding out low-performing teachers. Wysocki was fired from her job after receiving a low score based on IMPACT's algorithm, which promised to fairly evaluate teachers. But IMPACT was imperfect in that it couldn't measure the human factors that go into teaching and learning (such as struggles students might be facing at home), so O'Neil characterizes it as a "**weapon of math destruction**."

MINOR CHARACTERS

Kyle Behm – Kyle Behm is a young man who found that personality tests on job applications were keeping him from attaining employment at a large supermarket change. O'Neil asserts that the personality tests' algorithms were "**weapons of math destruction**."

Tim Clifford – Tim Clifford is a middle school English teacher in New York City. He was the victim of value-added modeling that assigned him a low teaching score due to faulty data.

Marc Leder – Marc Leder is a private equity investor who hosted a now-infamous fundraiser for 2012 U.S. presidential candidate Mitt Romney.

Rayid Ghani – Rayid Ghani is a data scientist who pioneered micro-targeting tools during Barack Obama's 2012 presidential campaign.

TERMS

Quant – A "quant," or quantitative analyst, is a specialist who uses mathematical and statistical methods to analyze the economy.

Hedge Fund – A hedge fund is a partnership of investors that uses high-risk methods to make a profit. Hedge funds focus on small fluctuations in the economy, train algorithms to predict errors and price swings, and place financial bets on those occurrences.

2007-2008 Financial Crisis – The 2007-2008 financial crisis was a worldwide economic downturn that began in the United States. The crisis stemmed from banks lending subprime mortgages (in other words, lending mortgages to people at a high risk of not being able to pay the loans back), and the resulting burst of an unsustainable housing market. The values of subprime mortgage-backed securities plummeted, financial institutions took severe damage, and several major banks folded. The crisis spread across the globe, creating a large recession and major crises in countries like Iceland and Greece.

Big Data – Big Data, or the so-called "Big Data economy," is a field that seeks to analyze incredibly large data sets through models and algorithms—some of which **O'Neil** classifies as "weapons of math destruction" because of how harmful they can be on a societal level. More colloquially, the idea of "big data" refers to our modern era, in which data about us—where we live, what we shop for, how much money we make, and more—is, in turn, used to control important aspects of our lives. Big Data can impact how we're targeted by advertisers and politicians; whether we can secure loans, credit, or insurance policies; and more.

Value-added Model – Value-added models seek to measure a teacher's effect on their students' achievement by predicting how students will score on an assessment. Teachers are then either rewarded or reprimanded based on the gap between the model's expectation and the students' actual scores.

Recidivism Models/LSI-R – The LSI-R (short for Level of Service Inventory-Revised) is a recidivism model that seeks to determine how likely a criminal is to repeat their offense after being released from prison. These models were built to try to make the U.S. justice system both fairer and more efficient. But because recidivism models rely on data about prior involvement with police (something that's statistically more likely in low-income or predominately non-white neighborhoods), **O'Neil** suggests that they are hampered by bias and racism. Models like the LSI-R, O'Neil argues, are "weapons of math destruction" because of the feedback loops they create and the biases that structure them.

PredPol – PredPol, short for Predictive Policing, is software pioneered by a California-based start-up that uses historical crime data to show where and when crimes are most likely to occur. PredPol was pioneered to cut down on crime by allowing police to identify and patrol hotspots. But according to **O'Neil**, PredPol could worsen the police's unequal treatment of white and non-white people by targeting poor and minority neighborhoods, where nuisance crimes like public drunkenness are more common.

Stop and Frisk – "Stop and frisk" is a New York City Police Department practice of stopping, detaining, questioning, and frisking or searching civilians on the street—especially in low-income, high-crime neighborhoods. In the early 2010s, the NYPD reported stopping and frisking over 684,000 New Yorkers in one year. Nine out of ten people subjected to stop and frisk during that time period were found innocent—and 87 percent of those targeted were Black or Latinx.

FICO Score – FICO scores are the standard credit-scoring system in the United States. The FICO model was created to evaluate the risk of a person defaulting on a loan. And because the model was designed to be transparent, fair, and backed by consistently updated data, **O'Neil** asserts that it is not a "weapon of math destruction." However, e-scores—unregulated models that aggregate even more data than the FICO model—use hidden metrics and faulty data that threaten FICO's hold as the industry standard.

E-scores/E-scoring – E-scores and e-scoring systems aggregate everything from zip codes to internet behavior to purchase history to create unregulated algorithms. These algorithms are then used to determine whether a person is worthy of things like credit cards, loans, or insurance policies.

Redlining – Redlining is an illegal practice used by bankers and insurance companies to delineate neighborhoods in which they refuse to invest or operate. Modern-day redlining is often fueled by harmful algorithms (what O'Neil calls "weapons of math destruction"). Bankers and insurers create their own ratings, or e-scores, to determine how worthy a person is of a loan or insurance policy based on proxy data about similar groups of people.

Microtargeting – Microtargeting is any kind of personalized advertising targeted at a specific person (or kind of person) based on data gathered from a person's internet history or demographics. Microtargeting is vast, hard to understand, and unregulated—so according to **O'Neil**, it qualifies as a "weapon of math destruction."

THEMES

In LitCharts literature guides, each theme gets its own color-coded icon. These icons make it easy to track where the themes occur most prominently throughout the work. If you don't have a color printer, you can still use the icons to track themes in black and white.

HUMANITY VS. TECHNOLOGY

In *Weapons of Math Destruction*, author Cathy O'Neil writes that humanity is in the midst of a "new revolution." "Big Data"—a field that uses huge swaths of data to make various industries more efficient or profitable—is rapidly changing the way society functions. And while O'Neil acknowledges that data collection and computer algorithms can be helpful in certain contexts, she also warns that behind the scenes, much of modern life is dictated by machines rather than people. Companies are increasingly using computer-based algorithms to interpret data and make important decisions (like who gets interviewed for a job or who can secure a loan). Moreover, computers are controlling our lives in potentially unfair and inhumane ways, since the algorithms they depend on can be inaccurate, biased, or otherwise flawed—and O'Neil suggests that this will continue unless humans play a more active role in data interpretation. In order to keep technology's influence over humans in check, corporations must implement internal regulations on how data is gathered, employ data scientists to make sure it's interpreted correctly, and ensure that developers are held accountable for the effects of the algorithms they create.

In recent decades, more and more sectors of the economy have begun to gather and use data in increasingly sophisticated ways—and some of this data has the potential to improve human lives. For instance, in the mid-1980s, *U.S. News & World Report* started releasing data-backed college rankings. Many colleges that were dissatisfied with their rankings took steps to improve their schools by fundraising, admitting a more diverse student body, and building better infrastructures for their campuses. Another example is standardized tests, which gather data about students' performances in American public schools. These tests have the potential to help teachers tailor their curriculum to their students' needs and to direct more funding to school districts that need additional resources. Lastly, O'Neil uses the example of trucking companies that have begun to more closely track and surveil their truckers' rigs. By installing cameras and GPS devices and monitoring how truckers are driving at different hours of the day, they can gather data about when their drivers might be struggling to stay awake—and thus prevent tragic or fatal accidents.

But although data collection can benefit humanity, relying too much on computer algorithms poses an existential and moral problem. In the book's conclusion, O'Neil suggests that algorithms and mathematical models that promise to make life more efficient by erasing human bias and error also end up erasing things that only humans can do: imagine, invent, and self-correct. "Compared to the human brain, machine learning isn't especially efficient," O'Neil writes. Machines can't differentiate between the truth and lies—they can only analyze data. This means that mathematical models and programs may be encoded with human bias—they are created by humans, after all. And if biased or otherwise faulty algorithms run on their own without being vetted or regulated by humans along the way, they may make connections based on flawed associations. As "error-ridden" as these computer systems may be, they play a huge role in determining some of the most important parts of modern-day life. For instance, many colleges use algorithms to screen applicants' personal information, and financial institutions use algorithms to decide who can or can't secure credit. Algorithms have an enormous amount of control over human lives—and yet they're not regulated or held to any kind of standard.

O'Neil argues that to make sure these programs and machines stay on track, direct human involvement at every stage of an algorithm's development and implementation is necessary. First, O'Neil suggests, data should be gathered in more transparent ways, so that the public understands when and why their personal information is being used. Second, the creators of algorithms and models need to take greater responsibility for their creations—O'Neil suggests that they should be made to take an oath similar to the Hippocratic Oath doctors take before practicing medicine, whose most famous promise is to "do no harm." Lastly, O'Neil suggests that in-house data scientists (as well as external advisors) need to vet these algorithms before they're put to use, and make sure that they're gathering and using data fairly. Together, all of these steps would help ensure that technology doesn't have undue control over humanity, and that the algorithms various industries use are helpful rather than harmful. It's up to humanity, O'Neil suggests, to recognize that we're not living in a "techno-utopia." Instead, we're at a very fragile moment in history that needs to be navigated delicately in order to put a stop to the flawed, harmful algorithms that O'Neil calls "**weapons of math destruction**."

DISCRIMINATION IN ALGORITHMS

In the early pages of *Weapons of Math Destruction*, data scientist and author Cathy O'Neil shares her experience working at a major hedge fund at the onset of the 2007-2008 global financial crisis. At this time, she began to feel troubled when she realized that human bias is written into the algorithms that are used to determine crucial things like job proficiency, recidivism rates for criminals, creditworthiness, and insurability—especially because algorithms are supposed to be fairer and more objective than human beings. These harmful algorithms (which O'Neil calls "**weapons of math destruction**," or WMDs) create unfair discrimination against women, racial and ethnic minorities, and low-income people. And O'Neil argues that unless developers, data scientists, and corporations actively work to purge sexism, racism, and classism from their algorithms, then women, racial minorities, and financially insecure people will be further victimized as time goes on.

WMDs create greater inequality when it comes to race and ethnicity. Algorithms haven't eliminated racial bias—they've just "camouflaged it with technology." Mathematical models like the LSI-R—a recidivism model that uses a lengthy questionnaire to determine whether an offender released from prison is likely to commit another crime—seem to be fair at first glance. But the questions it asks to gather data relate to whether an offender has friends or family members who've been arrested or incarcerated before, what kind of neighborhood the offender is planning on returning to, and what the offender's history with police is. These questions, in O'Neil's opinion, are leading and unfair. Due to redlining (a form of housing segregation) and other racist historical precedents, Black and Latinx people are often statistically more likely to live in low-income, high-crime neighborhoods. And because young Black and Latino men are disproportionately targeted by programs like stop and frisk, they're likelier to have a record of prior police involvement—even if they were innocent. So, these ostensibly fair models are actually encoded with racism. Models like the LSI-R penalize non-white people by failing to address the racial biases that underpin U.S. society, which only deepens the country's racial divide. Computer algorithms have the potential to help remedy the U.S.'s history of racism—but instead, many algorithms further perpetuate racial bias.

WMDs also deepen divides along the line of sex, preventing women from having fair shots at the same opportunities given to men. In the 1970s, the admissions office at St. George's Hospital Medical School in London used an algorithm to sort through the many applications they received for a limited number of positions. The algorithm promised to make the applications process fairer and more efficient. But in practice, it systematically rejected resumes from people whose names seemed to indicate that they were immigrants or racial minorities, as well as from women (who might become mothers, impacting their value as laborers). While this anecdote happened several decades ago, it is an example of how human biases often infiltrate even seemingly "objective" programs when the data used to train the programs is itself encoded with bias. When algorithms are trained to discriminate against a certain kind of individual, this can have devastating implications. Women and racial minorities have long had to fight to assert their value as workers, and algorithms like the one employed by St. George's threatened to make things even harder for female and non-white employees. By training machines to ignore information related to race, gender, or other categories, developers could ensure that their algorithms aren't discriminatory—but all too often, models reflect their makers' implicit biases.

Class in the U.S., too, is both exploited and cemented by harmful WMDs. To explain how WMDs play a role in perpetuating classism, O'Neil gives a hypothetical example of a working-class person who wants a fair rate on car insurance.

Insurance companies use algorithms to determine insurance rates—and the data that feeds these algorithms isn't always directly related to what kind of driver a person is. Someone might be a safe and skilled driver, which would entitle them to a reasonable rate. But insurers also take into account things like living in a low-income neighborhood where drunk drivers or car-jackings might be more common can become an unfair liability. Furthermore, people who live in low-income neighborhoods often commute to higher-income ones for work—and more time on the road means greater liabilities. So, a low-income person may receive a higher insurance rate not because of their driving record, but because of factors like location and driving time that they have little control over. Shift-scheduling is another realm that's been largely automated by models and algorithms—but these algorithms prize efficiency over fairness. For instance, while it might save time and money for the same employee to close a store one night and open it the next morning, this creates stress and sleep deprivation for the employee. This kind of automated scheduling prevents working-class people from setting aside family time and investing in things like education and recreation—so it keeps them from enrolling in night school or pursuing a hobby, deepening the class divide practically as well as emotionally.

Because WMDs are unfair to women, working-class people, and minorities, O'Neil argues, they must be dismantled—in other words, companies and organizations must identify where they're being used and begin regulating them. Otherwise, society will only grow more stratified, with disadvantaged groups entrapped by the very systems that claim to offer them a more equal social standing.

FAIRNESS VS. EFFICIENCY

One of the primary reasons behind the creation of algorithms that author Cathy O'Neil calls **WMDs**—"weapons of math destruction" that are widespread, harmful, and largely hidden from the public—was the desire to make various industries more equitable and efficient. In the late 20th and early 21st centuries, there was a sharp rise in the use of data to create mathematical models and algorithms that would help make schooling, credit scoring, and even criminal sentencing both easier and more just. But over time, in many cases, being efficient won out over being fair—and O'Neil suggests that today, the U.S. (and indeed the whole world) has been saddled with algorithms and systems that prioritize speed, ease, and profitability over fairness and equity. In order to reform WMDs, O'Neil asserts, data scientists and tech companies alike must begin sacrificing efficiency and profits for the sake of fairness, transparency, and morality.

Algorithms were initially created to be both more fair and more efficient than humans ever could be. To illustrate why mathematical models became important in restoring fairness to the economy, O'Neil offers a general example of a banker offering a newlywed couple in a small town a load in the 1960s. This banker might have conscious or unconscious biases against the couple. If he didn't like their families, if they were a different race, or if he had some other prejudice against them, he could deny them a loan—even if they qualified for one. So, mathematical models were introduced to banking to remove these human biases from the equation. By quickly and efficiently determining who was creditworthy based on objective criteria, algorithms could create a fairer world.

Sixty years later, though, efficiency has won out over fairness—in O'Neil's words, "the world is dominated by automatic systems." O'Neil writes that "today, the success of a model is often measured in terms of profit [or] efficiency. But she questions whether these models are in fact actually "successful"—and whether society should redefine its idea of success. For example, advertisements that target internet users who are searching for information about food stamps sometimes collect information about these users, then use that information to target them with even more predatory ads—like ones promoting for-profit colleges. And recidivism models like the LSI-R are ostensibly efficient in determining which criminals are most likely to become repeat offenders. But by making huge determinations about people's freedom based on data rather than a person's humanity, these models also prioritize efficiency over fairness. When these algorithms rake in money or otherwise produce favorable results for companies and organizations, they're considered a success—even though they take advantage of disenfranchised, vulnerable people in order to "succeed." O'Neil suggests that when success is tied to efficiency and profit in this way, it means that fairness and equity aren't part of the metrics of a successful algorithm. By preying on the poor and failing to prioritize justice and objectivity, algorithms are deepening inequality around the world.

Efficiency and profit, then, shouldn't be the metrics of a successful model—fairness and equity should, and so companies must start to use their algorithms for good. There are models, O'Neil shows, that are already seeking to do good in the world. For example, Mira Bernstein, a Harvard PhD in mathematics, created a model that would scan industrial supply chains to look for signs of forced labor (or modern-day slavery) for a non-profit working to root out slave labor in the global economy. But models like this one, O'Neil states, are not prevalent enough. Algorithms and models are more frequently used to make the most profits in the most efficient way for big companies—models that prioritize fairness and that help nonprofits or other social justice initiatives reach their goals simply aren't as valuable in the Big Data economy (the field that gathers and analyzes large sets of data). While algorithms were initially created to remove humanity from the equation and combine fairness and efficiency, the automatic nature of mathematical modeling has tilted toward efficiency over fairness. Now, O'Neil asserts, it's time to put the humanity back into the picture rather than leaving the issue to the marketplace—which will, she predicts, always prize "efficiency, growth, and cash flow."

DATA, TRANSPARENCY, AND U.S. DEMOCRACY

One of the hallmark qualities of a **WMD**—a "weapon of math destruction," or a destructive mathematical algorithm—is, in author Cathy O'Neil's view, the fact that it's "opaque." In other words, the systems that govern it (and sometimes its overall purpose) are kept secret or shrouded in mystery. Things like FICO credit scores and baseball statistics used in game wagers are transparent: anyone can access them. But tech companies like Google and Facebook use decidedly opaque methods to gather data from their users, while companies like Sense Networks are transparent in how they gather data but completely mum as to how they're using it. This tendency toward secrecy, O'Neil asserts, poses a threat not just to individual citizens' privacy, but to the fabric of American society as whole. Data collection is playing an increasingly important role in U.S. civic life (it influences political polling, campaign advertising, public services like housing assistance, and voting). Therefore, O'Neil argues, its misuse threatens the transparency and legitimacy of public institutions and democratic processes like federal aid, major elections, and data-backed public policy decisions.

Big tech companies have set a precedent for using hidden methods to gather, interpret, and use data. Targeted ads are now a part of daily life. Based on how people navigate the internet, shop and bank online, and more, they'll be shown certain kinds of advertisements that seek to gather more of their data and personal information. Then, large companies can analyze or sell off people's data, and influence consumers to buy goods and services based on their browsing history. But companies like Facebook aren't just using user data to sell products—they're using it to determine what kinds of news people should see, how they should interpret it, and how it will affect them. In the run-up to the 2016 U.S. presidential election, Facebook came under fire for manipulating its users' news feeds to gather data on how they'd interact with different kinds of news. But Facebook's actions prioritized gathering and analyzing data about user behavior—not ensuring that its users got accurate, fairly reported news. Data scientists at Facebook were tampering with people's feeds, and in the process, they were contributing to a dangerous rise of misinformation—especially misinformation concerning the turbulent political climate and "fake news" about each presidential candidate.

Now, political campaigns and government institutions in the U.S. are using big tech's same methods of gathering and using data in a manipulative way. In 2011, Rayid Ghani—a data scientist for Barack Obama's reelection campaign—used software he'd pioneered as an analyst at the consulting firm Accenture to gather data on swing voters (people who aren't firm supporters of any one candidate or political party). Then, he used the software to find information about millions of Americans who fit the profiles of those swing voters and targeted them all with political advertisements. This process is called "microtargeting," and it is useful to campaigns—but in most scenarios, it's also an invasion of personal privacy. On its own, targeting potential voters with ads isn't unprecedented or even necessarily harmful. But by keeping tabs on users' "likes" and using those likes to rank people's assumed personality traits, political efforts like the Obama campaign (and, later, Hillary Clinton's 2016 presidential campaign) are essentially turning the voting public into a financial market.

City and state governments, too, are increasingly using data to determine how they function. In 2013, O'Neil worked at New York City's Housing and Human Services Departments for a time, building a model that would help get houseless people out of shelters and into homes. But the city didn't want to spend the money on Section 8 housing vouchers, even though O'Neil's data proved that the vouchers helped disadvantaged people. They ignored O'Neil's research and instead poured their resources into a new program that would limit housing subsidies significantly. In this case, the city government ignored how data could help people, focusing only on how microtargeting ads for a predatory program could keep budgets low.

When government institutions start acting like private tech companies, O'Neil asserts, American democracy comes under threat. Tactics like microtargeting and reckless, mass data-gathering "infect our civic life." Flawed, faulty data now dictates how politicians campaign and how newscasters report on political happenings. A good example of this is Hillary Clinton's 2016 presidential campaign: she was, according to most polls and mainstream media outlets in the weeks leading up to the election, the predicted winner. But she ended up losing to Donald Trump, which suggests that data can't always be trusted. And when lobbyists and interest groups can target American citizens with misinformation—dangerously false or misleading facts that can sway how people vote—the integrity of democracy is threatened. By influencing people's political opinions with flawed data, groups that should be pursuing transparency and democracy instead favor profits and fast, easy solutions.

When people don't have all the information, they can't make good decisions or faithfully perform their civic duties. Though algorithms, microtargeting, and misinformation are all here to stay, O'Neil suggests that "the [U.S.] government [...] has a powerful regulatory role to play." In Europe, any data that's gathered and collected must be approved by the user—and O'Neil suggests that similar measures in the U.S. could help make sure that transparency and user autonomy are protected. Government policy change and corporate accountability are necessary, O'Neil states, to make sure that the public is protected from the ongoing spread of misinformation and increasingly sneaky tactics that invade citizens' privacy. When the public is shut out of decision-making about how their information is gathered, interpreted, and used, O'Neil suggests, their freedom is profoundly threatened.

SYMBOLS

Symbols appear in **teal text** throughout the Summary and Analysis sections of this LitChart.

WEAPONS OF MATH DESTRUCTION

The titular term "weapons of math destruction" represents the serious harm that certain kinds of algorithms can cause to global society. A "weapon of math destruction," or WMD for short, is a term coined by Cathy O'Neil to describe a dangerous mathematical algorithm or model. There are a few core characteristics of a WMD: it must be opaque (meaning its methods of gathering or using data are purposefully hard to ascertain), it must be widespread, and it must be damaging.

Throughout the book, O'Neil repeatedly compares the algorithms that various organizations use to gather information about people to weapons of mass destruction. Much like a weapon of *mass* destruction—a nuclear bomb, for instance—a weapon of *math* destruction misuses math to cause widespread damage. This is because data-driven algorithms are often encoded with human bias and can therefore cause damage by preying on people (through targeted political ads, for instance) or discriminating against them (by automatically denying them opportunities based on characteristics like sex, race, or class).

With this comparison, O'Neil characterizes the Big Data economy as a kind of war zone and suggests that the algorithms that govern it are indeed deadly weapons. Through the idea of WMDs, O'Neil underscores how potentially harmful these weapons of math destruction are to individual lives as well as society, as they influence major aspects of daily life, interfere with politics and democracy, and deepen social divides.

QUOTES

Note: all page numbers for the quotes below refer to the Broadway edition of *Weapons of Math Destruction* published in 2017.

Introduction Quotes

The math-powered applications powering the data economy were based on choices made by fallible human beings. Some of these choices were no doubt made with the best intentions. Nevertheless, many of these models encoded human prejudice, misunderstanding, and bias into the software systems that increasingly managed our lives. Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists. Their verdicts, even when wrong or harmful, were beyond dispute or appeal. And they tended to punish the poor and the oppressed in our society, while making the rich richer.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 3

Explanation and Analysis

In this passage, author Cathy O'Neil explains the logic behind the term "weapons of math destruction"—harmful models that aggregate and interpret data. The "data economy," or "Big Data," is the sector of the economy that revolves around the harvesting, sorting, and interpreting of people's personal data. Models and algorithms "made by fallible human beings" power this economy. So, as these hidden, complex, and indeed biased systems write "verdicts" about what kind of shoppers, citizens, and human beings we are, they threaten everything from people's ability to secure a loan to the very foundations of U.S. democracy. By victimizing the poor while "making the rich richer," these models make global society more stratified, more unfair, and more unstable.

This passage explains what WMDs are, how they work, and why they're dangerous. Models that interpret data have become modern day "gods"—they're seen as mysterious yet fair and unimpeachable. But they're riddled with errors that destabilize major societal structures (like education systems and labor networks). So, by failing to question, regulate, or vet WMDs, humanity has allowed the Big Data economy to take hold over huge parts of everyday modern life. As the book goes on, O'Neil will show how WMDs now control everything from elementary school classrooms to college admissions to credit checks to insurance—and she'll explain why it's dangerous for unregulated algorithms to control sensitive issues tied to race, class, and politics.

Do you see the paradox? An algorithm processes a slew of statistics and comes up with a probability that a certain person might be a bad hire, a risky borrower, a terrorist, or a miserable teacher. That probability is distilled into a score, which can turn someone's life upside down. And yet when the person fights back, "suggestive" countervailing evidence simply won't cut it. The case must be ironclad. The human victims of WMDs, we'll see time and again, are held to a far higher standard of evidence than the algorithms themselves.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 10

Explanation and Analysis

Here, Cathy O'Neil describes the crude (and cruel) nature of WMDs in order to show how they diminish people's agency and value. WMDs are everywhere—they're used in college admissions, in job-seeking, and in processes like securing loans and getting covered under auto insurance. They quickly (but sometimes incorrectly) process and interpret data about an individual in order to predict what kind of person they are. In other words, WMDs are treating humans like products. What's dangerous about WMDs is that they hold people to impossibly high standards, weeding out people with less-than-stellar credit scores who apply for loans or rejecting a person's job application based on their answers to a digital personality test.

Yet WMDs themselves are rarely vetted, regulated, or updated—they're not held to very high standards at all. This is dangerous, in O'Neil's estimation, because algorithms and models are supposed to be used to make modern life both fairer and more efficient. Yet they're decidedly unfair, and they're not always efficient, because they often rely on faulty or flawed data. So, she suggests that WMDs need to be transparent, regulated, and held to higher standards.

Chapter 1: Bomb Parts Quotes

The value-added model in Washington, D.C., schools [...] evaluates teachers largely on the basis of students' test scores, while ignoring how much the teachers engage the students, work on specific skills, deal with classroom management, or help students with personal and family problems. It's overly simple, sacrificing accuracy and insight for efficiency. Yet from the administrators' perspective it provides an effective tool to ferret out hundreds of apparently underperforming teachers, even at the risk of misreading some of them.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 21

Explanation and Analysis

Here, author Cathy O'Neil delves more deeply into why many "weapons of math destruction" (WMDs) are inefficient, even while they're touted as solutions to making modern life more efficient and streamlined. Value-added models—models that seek to measure the gap between students' targeted improvement and their actual improvement from year to year—are, in O'Neil's estimation, WMDs. This is because they claim to be fair and efficient, yet they're oversimplistic to the point of being irresponsible.

These models use data in the form of student test scores to evaluate teachers' performances, but they're so rudimentary that they're not reliable tools in determining whether teachers are doing their jobs well. The data these models measure doesn't take into account the human aspect of the student-teacher relationship. So, while these models might be efficient in terms of sorting test scores, they shouldn't be used to make decisions about teachers' futures.

But because WMDs are generally created to maximize efficiency and profit, these models are indeed doing the jobs they were designed to do. They're unfair, though they claim to be fair—and they're not reliable, though they claim to be revolutionizing the efficiency of an entire sector of society (the education system). They're damaging and destructive, and they must be reformed and regulated in order to truly look out for teachers and students across the U.S.

And here's one more thing about algorithms: they can leap from one field to the next, and they often do. Research in epidemiology can hold insights for box office predictions; spam filters are being retooled to identify the AIDS virus. This is true of WMDs as well. So if mathematical models in prisons appear to succeed at their job—which really boils down to efficient management of people—they could spread into the rest of the economy along with the other WMDs, leaving us as collateral damage.

That's my point. This menace is rising.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 31

Explanation and Analysis

In this passage, author Cathy O'Neil explains why WMDs represent such a major threat to society. WMDs, or weapons of math destruction, is the term O'Neil ascribes to models and algorithms that are deceitful or hidden, widespread, and damaging to society. For instance, recidivism models that attempt to predict whether a criminal will become a repeat offender are WMDs because they interpret data that's not gathered transparently, they're widespread, and they can cause real harm to people.

In O'Neil's estimation, WMDs are so dangerous because they can "spread into the rest of the economy." Already, WMDs can be found in the college admissions process, the job application process, and the policing system, and many other realms of society. These algorithms use faulty data and oversimplistic interpretation of that data to make snap decisions in the name of efficiency—and they're warping society in the process. By regulating WMDs to make sure that they're both efficient and fair, humanity can avoid the creation of a society "menace[d]" by computer programs that see people as things to "manage" rather than full, complex human beings.

Chapter 2: Shell Shocked Quotes

Paradoxically, the supposedly powerful algorithms that created the market, the ones that analyzed the risk in tranches of debt and sorted them into securities, turned out to be useless when it came time to clean up the mess and calculate what all the paper was actually worth. The math could multiply the horseshit, but it could not decipher it. This was a job for human beings. Only people could sift through the mortgages, picking out the false promises and wishful thinking and putting real dollar values on the loans.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 43

Explanation and Analysis

In this passage, O'Neil describes how humanity came to the rescue when technology failed during the 2007-2008 financial crisis. After Wall Street slipped into chaos in 2007, setting off a chain reaction that would affect the entire global economy, it became clear that fudged and faulty data had ground the market to a halt. But the computer programs that had facilitated the disaster couldn't be trusted to "decipher" it—only humans could sift through the mess and begin repairing the economy from the inside out. O'Neil believes that this is important because it suggests that in the battle between humanity and technology, even the most sophisticated technology (and the advanced models that support it) cannot keep up with human reasoning.

This illustrates the fact that "weapons of math destruction" aren't just unreliable—they're actively threatening the health of human society. And yet humans will inevitably have to pick up the pieces when the so-called "Big Data economy" fails us. This illustrates the need for a human hand to be placed back into the world of tech. Humans need to vet, examine, and regulate the programs that increasingly control our lives—or we will be left sifting through the ashes of another chaotic explosion caused by these metaphorical weapons.

Chapter 3: Arms Race Quotes

What does a single national diet have to do with WMDs? Scale. A formula, whether it's a diet or a tax code, might be perfectly innocuous in theory. But if it grows to become a national or global standard, it creates its own distorted and dystopian economy. This is what has happened in higher education.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 51

Explanation and Analysis

Here, author Cathy O'Neil introduces the scourge of "weapons of math destruction" (WMDs) in higher education by describing the chaos that would erupt if the U.S. was made to follow a hypothetical standardized diet. O'Neil uses the metaphor of a strict, regulated diet becoming the standard across the U.S.—and plunging the economy into free-fall as a result—in order to show that the scale and influence of WMDs are indeed dangerous.

Faulty algorithms that incorrectly interpret flawed data are at the core of many major economic sectors—including higher education. Algorithms now sort through prospective students' applications—and they are increasingly overlooking students' humanity in favor of easily interpreted data like test scores. By flattening out the admissions process, these WMDs are stripping applicants of their humanity and reducing them to numbers. If this process continues to spread unchecked throughout global society, humanity will find itself in a "distorted and dystopian" world.

It sounds like a joke, but they were absolutely serious. The stakes for the students were sky high. As they saw it, they faced a chance either to pursue an elite education and a prosperous career or to stay stuck in their provincial city, a relative backwater. And whether or not it was the case, they had the perception that others were cheating. So preventing the students in Zhongxiang from cheating was unfair. In a system in which cheating is the norm, following the rules amounts to a handicap.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 63

Explanation and Analysis

In this quotation, O'Neil describes how a group of students in rural China protested to assert their right to cheat on school entrance exams. They did this so that they'd be on a level playing field with more well-off students from the country's privileged major cities. This anecdote is important to O'Neil because it illustrates how "weapons of math destruction" (WMDs) are warping norms and destabilizing society.

There's an enormous divide between the rich and the poor in countries like China. So, for students in rural areas who are aspiring to better futures, the only way to level the playing field (without access to specialized tutoring programs and high-caliber schooling systems, not to mention insider information about the tests available to wealthy students) is to try to get a leg up by cheating. These students know that their admission to good schools—and thus their ability to secure good jobs and good lives—is completely contingent on data (test scores) that's harvested and interpreted by WMDs (admissions algorithms). In the name of efficiency, WMDs all over the world are making society more and more stratified and preventing working-class people from changing their station in life.

Chapter 4: Propaganda Machine Quotes

The Internet provides advertisers with the greatest laboratory ever for consumer research and lead generation. [...] Within hours [...], each campaign can zero in on the most effective messages and come closer to reaching the glittering promise of all advertising: to reach a prospect at the right time, and with precisely the best message to trigger a decision, and thus succeed in hauling in another paying customer. This fine-tuning never stops.

And increasingly, the data-crunching machines are sifting through our data on their own, searching for our habits and hopes, fears and desires.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 75

Explanation and Analysis

In this passage, Cathy O'Neil describes how online advertisers use algorithms that qualify as "weapons of math destruction" to gather data about people—and use it to get them to buy things. These algorithms, O'Neil asserts, are predatory and dangerous. While they may seem relatively harmless—after all, they're just suggesting people buy things—they actually represent a much more dangerous trend in how technology and data inform our lives as human beings. These algorithms don't just track our desires—they learn what we're afraid of. They know our financial standing, so they can advertise predatory loans. They also know where we live, and from our zip codes, they target us based on our presumed race and income. As a result, advertisers have unlimited access to information about us—and they try to use it to maximize their profits without considering how this process might affect the people they're targeting.

O'Neil uses this passage to show her readers that predatory ads are the results of predatory algorithms that use the facts of our own lives to manipulate us. This example suggests that while ads (and the algorithms that fuel them) may offer "glittering promise[s]," they're actually preying on vulnerable people by targeting them based on their "hopes, fears and desires."

For-profit colleges, sadly, are hardly alone in deploying predatory ads. They have plenty of company. If you just think about where people are hurting, or desperate, you'll find advertisers wielding their predatory models.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 81

Explanation and Analysis

Here, Cathy O'Neil further explores how "weapons of math destruction" (WMDs)—faulty, complex, harmful algorithms—fuel predatory ads. Many targeted online advertisements are fairly harmless—they encourage us to buy things we've already looked for online, or they show us a product related to something we've recently purchased. But many online ads are predatory—and WMDs aim them at people who are "hurting" or "desperate" to improve their lives. For instance, a person who's recently sought out information about food stamps online might get a targeted ad for a predatory for-profit college that promises to help them get a quick degree and thus a better job, while someone with student debt might get an ad for a predatory loan refinancing company.

By appealing to people who are desperate and in dire straits, these WMDs deepen social divides and keep the poor struggling (while the rich get richer). Rather than serving people in need with ads for programs that could help them, WMDs are used by companies looking to maximize their profits as much as possible, without a second thought for who's hurt in the process. This, O'Neil asserts, is why WMDs are so dangerous—and why they'll only become more harmful as they spread across different sectors of the economy.

Chapter 5: Civilian Casualties Quotes

These types of low-level crimes populate their models with more and more dots, and the models send the cops back to the same neighborhood.

This creates a pernicious feedback loop. The policing itself spawns new data, which justifies more policing. And our prisons fill up with hundreds of thousands of people found guilty of victimless crimes. Most of them come from impoverished neighborhoods, and most are black or Hispanic. So even if a model is color blind, the result of it is anything but. In our largely segregated cities, geography is a highly effective proxy for race.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 87

Explanation and Analysis

Here, Cathy O'Neil describes the insidious feedback loops that are created when "weapons of math destruction" (WMDs) play a role in policing in the U.S. In recent years, police departments around the country have adopted PredPol, a predictive policing software that attempts to forecast where and when crimes are most likely to happen based on huge swaths of data about different neighborhoods and historic crime rates. While PredPol seems like it would make society fairer and safer by cutting down on crime, in reality, the software is fundamentally flawed.

As O'Neil describes here, PredPol essentially affirms its own biases and outcomes—and this creates a more dangerous environment for Black and Latinx people (as well as people from working-class neighborhoods or other ethnic or racial minority groups). Because police officers are agents of the state, this targeting of historically oppressed groups threatens the very fabric of U.S. society. WMDs prioritize efficiency over fairness—and if police forces increasingly do the same, the already vulnerable people will face even greater risks.

> Police make choices about where they direct their attention. Today they focus almost exclusively on the poor. [...] And now data scientists are stitching this status quo of the social order into models, like PredPol, that hold ever-greater sway over our lives.
>
> The result is that while PredPol delivers a perfectly useful and even high-minded software tool, it is also a do-it-yourself WMD. In this sense, PredPol, even with the best of intentions, empowers police departments to zero in on the poor, stopping more of them, arresting a portion of those, and sending a subgroup to prison. [...]
>
> The result is that we criminalize poverty, believing all the while that our tools are not only scientific but fair.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 91

Explanation and Analysis

In this passage, author Cathy O'Neil describes how PredPol, a predictive policing software, deepens social divisions by criminalizing poverty. She calls PredPol a "do-it-yourself WMD," asserting that it's only a matter of time before its impact is widespread and dangerous. This is because nuisance crimes (like public intoxication or vandalism) happen more regularly in low-income neighborhoods. But crime is more widespread—and more complex—than predictive policing programs can process. White collar crimes, like tax fraud and money laundering, are crimes committed by people who can afford to get away with them, because the police aren't necessarily looking for those kinds of crimes. And because the police aren't looking for them, they're not discovering them—so there isn't data to feed the software that predicts criminal activity hotspots.

So essentially, PredPol and other programs like it target the poor while letting the rich get away with crimes that are arguably more serious and damaging to society. O'Neil uses this anecdote to show that while these predictive policing programs might be efficient, they certainly aren't fair—and they're threatening to make society in the U.S. even more stratified. As agents of the state, police officers have a duty to be arbiters of justice, and yet they're actually perpetuating inequality.

> While looking at WMDs, we're often faced with a choice between fairness and efficacy. Our legal traditions lean strongly toward fairness. The Constitution, for example, presumes innocence and is engineered to value it. [...]
>
> WMDs, by contrast, tend to favor efficiency. By their very nature, they feed on data that can be measured and counted. But fairness is squishy and hard to quantify. It is a concept.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 94-95

Explanation and Analysis

Here, Cathy O'Neil describes the "choice between fairness and efficacy" that lies at the heart of every "weapon of math destruction." Companies originally implemented algorithms to make themselves fairer and more efficient—to remove the human element from processes that were full of potential bias (like in the case of a Black family applying for a mortgage and facing down a racist banker in the process). But over time, these models have begun sacrificing fairness in the name of efficiency. Now, when people apply for mortgages, they might not have to face down a person who'd judge them, but they're vulnerable to algorithms that determine whether they're credit-worthy based on irrelevant data like the zip code in which they live.

And yet the foundation of the U.S. Constitution is, ostensibly, fairness. The justice system, for example, presumes innocence—so it sacrifices efficiency in the name of fairness when it releases a potential criminal because there isn't enough evidence to convict them. But WMDs are just the opposite—as they proliferate throughout society, they're essentially asserting that fairness isn't a worthy metric. And, again, because the U.S.'s legal and justice systems are rooted in the prioritization of fairness, these WMDs are quite literally changing and threatening the fabric of U.S. society.

Chapter 6: Ineligible to Serve Quotes

❝❝ The hiring business is automating, and many of the new programs include personality tests like the one Kyle Behm took. It is now a $500 million annual business and is growing by 10 to 15 percent a year […]. Such tests now are used on 60 to 70 percent of prospective workers in the United States […].

Naturally, these hiring programs can't incorporate information about how the candidate would actually perform at the company. That's in the future, and therefore unknown. So like many other Big Data programs, they settle for proxies. And as we've seen, proxies are bound to be inexact and often unfair.

Related Characters: Cathy O'Neil (speaker), Kyle Behm

Related Themes:

Related Symbols:

Page Number: 108

Explanation and Analysis

Here, Cathy O'Neil describes how "weapons of math destruction" (WMDs) have infiltrated the hiring industry. WMDs are dangerous, no matter what economic sector in which they appear. The employment of faulty algorithms that use proxy data—for example, the answers to a personality test rather than the work experience gathered from an applicant's resume—is transforming the hiring industry. And as the industry becomes a site of enormous profits, as O'Neil shows it is here, it's prioritizing efficiency and easy answers over fairness.

These algorithms claim to be able to predict who will be a good employee and who won't. But in the case of someone like Kyle Behm, a student who applied for a grocery store job after a year out of school dealing with mental health issues, the algorithms border on illegal as they use irrelevant personal information to try to predict a person's future. Like all WMDs, the WMDs used in the hiring sector are, in O'Neil's estimation, extremely dangerous. They're preventing working-class people, people with criminal records, and people with a history of chronic illness from having a fair shot at a better future—and they're doing so in the name of profit.

❝❝ The key is to analyze the skills each candidate brings […], not to fudge him or her by comparison with people who seem similar. What's more, a bit of creative thinking at St. George's could have addressed the challenges facing women and foreigners. […]

This is a point I'll be returning to in future chapters: we've seen time and again that mathematical models can sift through data to locate people who are likely to face great challenges, whether from crime, poverty, or education. It's up to society whether to use that intelligence to reject and punish them—or to reach out to them with the resources they need. We can use the scale and efficiency that make WMDs so pernicious in order to help people.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 117

Explanation and Analysis

In this passage, Cathy O'Neil continues her investigation of WMDs in the hiring sector by sharing an anecdote about how St. George's, a hospital in London. The hospital got into trouble in the 1970s by using software that tossed out applications from women and people of color to narrow down applicants for available jobs. The example of St. George's prejudicial hiring process is important because it shows that many of these algorithms exist only to perpetuate racism, sexism, and classism. The hiring managers at St. George's felt that people who spoke English as a second language and women (who could become pregnant and leave work for long stretches of time) were liabilities as new hires. So, they simply excluded these people's resumes from consideration.

What O'Neil points out in this passage is that WMDs like the one St. George's used could be used for the opposite purpose: to single out female and minority applicants and connect them with opportunities that would help them grow. Non-native English speakers could've been hired and put in touch with language resources or courses that would help them gain full proficiency in a second language—no doubt an asset to the hospital rather than a liability. And female applicants could've been shown support for whatever life choices they did or didn't make—life choices that an algorithm never could have predicted. WMDs are efficient, and they are increasingly prominent in spheres like the hiring sector. So, through this example, O'Neil is suggesting that their widespread influence should be used to accommodate rather than exclude people.

Phrenology was a model that relied on pseudoscientific nonsense to make authoritative pronouncements, and for decades it went untested. Big Data can fall into the same trap. Models like the ones that red-lighted Kyle Behm and blackballed foreign medical students at St. George's can lock people out, even when the "science" inside them is little more than a bundle of untested assumptions.

Related Characters: Cathy O'Neil (speaker), Kyle Behm

Related Themes:

Related Symbols:

Page Number: 117

Explanation and Analysis

Here, Cathy O'Neil compares Big Data's ostensibly predictive powers to the racist and long-debunked study of phrenology. This was a pseudo-scientific field that claimed to be able to predict things about certain people (and entire races) by using bumps, dips, and other irregularities of the skull to determine personality traits and future behavior. In this way, phrenology was racist, sexist, and classist—and it was completely unfounded in science.

Similarly, "weapons of math destruction" (WMDs) that are used throughout the Big Data economy aren't rooted in facts or science. Yet they predict how a person might behave as a student or employee or what liabilities they might incur as a homeowner, credit-holder—and they often use race, class, and gender as determining factors. Thus, O'Neil uses this comparison between phrenology and WMDs to show just how dangerous WMDs are. They're being used to justify the mistreatment and sidelining of entire swaths of people around the world—and this is directly contributing to greater social stratification. WMDs need to be seen for what they are: inaccurate and dangerous models that aren't truly able to reliably predict anything.

Chapter 7: Sweating Bullets Quotes

With Big Data, [...] businesses can now analyze customer traffic to calculate exactly how many employees they will need each hour of the day. The goal, of course, is to spend as little money as possible, which means keeping staffing at the bare minimum while making sure that reinforcements are on hand for the busy times.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 124

Explanation and Analysis

In this passage, Cathy O'Neil describes how huge swaths of data are being used to optimize businesses around the world with the goal of maximizing profit—but without any consideration for employees' well-being. It isn't just the hiring sector that's been overtaken by "weapons of math destruction" (WMDs)—on-the-job scheduling, too, is now automated by algorithms. These models look at how a business is scheduling their employees, what hours are busiest each day of the week, and when things are slow—and they create scheduling around that data. This often means that employees' schedules are erratic and demanding. Because an employer's profits are prioritized over their employees' scheduling needs, this means that working-class people must build their lives around the data that's controlling them—leaving them with little room for their own pursuits.

Again, O'Neil is using this passage to show how by prioritizing efficiency over fairness, WMDs—and the businesses that use them—are keeping working-class people struggling while allowing the rich to get richer. Society is becoming more stratified and class divisions are growing deeper as a direct result of these WMDs' influence.

> But data studies that track employees' behavior can also be used to cull a workforce. As the 2008 recession ripped through the economy, HR officials in the tech sector started to look at those Cataphora charts with a new purpose. They saw that some workers were represented as big dark circles, while others were smaller and dimmer. If they had to lay off workers, and most companies did, it made sense to start with the small and dim ones on the chart.
>
> Were those workers really expendable? Again we come to digital phrenology. If a system designates a worker as a low idea generator or weak connector, that verdict becomes its own truth. That's her score.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 132

Explanation and Analysis

This passage explains how a company called Cataphora developed software that would allow them to harvest data from employee emails to determine which employees were, based on their interoffice communications, most engaged and most useful. Cataphora's "weapon of math destruction" (WMD) was particularly destructive, O'Neil shows, because it used unreliable and ultimately irrelevant data—how employees were communicated via email and instant messaging—to determine which employees were ideas generators, and which employees weren't as communicative or innovative. The employees that weren't seen as being useful enough were declared expendable, and some were terminated.

The idea of Cataphora's software expanding throughout workplaces around the world is frightening and disturbing to O'Neil because it's more of the "digital phrenology" (pseudoscience) that claims to use data and science to predict the future—when really, such a thing is impossible. People are being reduced to data harvested from their emails, and their actual job performances are no longer the most important factor in whether their company asks them to stay or leave. By reducing people to a "score," WMDs are dehumanizing people and allowing companies to focus on hyper-productivity and maximization of profits over any other consideration.

> While its scores are meaningless, the impact of value-added modeling is pervasive and nefarious. "I've seen some great teachers convince themselves that they were mediocre at best based on those scores," Clifford said. "It moved them away from the great lessons they used to teach, toward increasing test prep. To a young teacher, a poor value-added score is punishing, and a good one may lead to a false sense of accomplishment that has not been earned."

Related Characters: Tim Clifford, Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 139

Explanation and Analysis

In this passage, O'Neil speaks with Tim Clifford—a New York City teacher who nearly lost his job as a result of bogus scores resulting from a value-added model's assessment of him as an educator. In Clifford's case, a value-added model attempted to measure the gap between how the algorithm thought Clifford's students should be scoring and how they actually were scoring. This nearly ended his teaching career, because he scored very low on his evaluation. But when he realized that other talented teachers at his school were also scoring unbelievably low, he began to realize that the model was faulty.

O'Neil and Clifford believe that value-added modeling is "pervasive and nefarious" because it's destroying teachers' confidence, while encouraging them to see their students as numbers or scores rather than as full human beings. Especially for teachers of young students, there's a lot more that goes into providing a child with a safe space and a good education than test scores—but these models, which seek to weed out inefficient teachers who aren't constantly raising their classes' scores year to year, are turning students into cogs in the machine of the Big Data economy. This is dangerous because it starts children's lives off by telling them that they're only as good as their test scores—and that if they're seeking to improve their lives through good schooling, they need to meet certain arbitrary metrics.

Chapter 8: Collateral Damage Quotes

Since [the invention of the FICO score], the use of scoring has of course proliferated wildly. Today we're added up in every conceivable way as statisticians and mathematicians patch together a mishmash of data, from our zip codes and Internet surfing patterns to our recent purchases. Many of their pseudoscientific models attempt to predict our creditworthiness, giving each of us so-called e-scores. These numbers, which we rarely see, open doors for some of us, while slamming them in the face of others. Unlike the FICO scores they resemble, e-scores are arbitrary, unaccountable, unregulated, and often unfair—in short, they're WMDs.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 143

Explanation and Analysis

In this passage, O'Neil describes the dangers of a new kind of "weapon of math destruction" (WMD)—e-scores, or models that aggregate random data about a person's life in order to determine whether they're credit-worthy or not. E-scores are a WMD that have spiraled out of control after ostensibly noble beginnings. In the past, people would have to visit their local bankers in person to get a loan, a credit card, or a mortgage—and that meant that the banker could judge them based on personal information he knew about them, the color of their skin, or other biased factors.

The invention of the FICO score, a vetted data aggregation that took a holistic look at a person's credit, promised to make the process of securing credit more objective and efficient. But as banks and lending companies began creating their own unregulated versions of FICO scoring known as e-scores, the process once again swung back toward the judgmental and unfair.

Now, e-scoring uses proxy data like the zip code in which a person lives to determine what kind of credit-holder or loan recipient they're likely to be. And as a result, e-scores are vulnerable to prejudice and inaccuracy. So now, the pendulum has swung too far in the opposite direction—these sophisticated yet inherently flawed WMDs make huge decisions about people's futures, and there's no way to appeal to them and show them who an individual truly is. As e-scores become the norm, O'Neil worries, the world will become even more unfair—and it will be impossible for non-white people, women, and working-class people to prove their humanity and their worth to these unfeeling (and deeply flawed) algorithms.

Chapter 9: No Safe Zone Quotes

So why would [auto insurance companies'] models zero in on credit scores? Well, like other WMDs, automatic systems can plow through credit scores with great efficiency and at enormous scale. But I would argue that the chief reason has to do with profits.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 165

Explanation and Analysis

In this passage, Cathy O'Neil explains how auto insurance companies' collective desire to maximize profit is forever changing how people get insured—and whether they can get insured at all. When determining who is eligible for auto insurance—something every driver needs—and what their premiums will look like, insurance companies increasingly use proxy data like credit scores and the zip codes in which people live to decide these things. It would make more sense, of course, for these companies to look at people's driving records—but even a DUI, O'Neil argues, counts less nowadays than a person's credit score when it comes to car insurance. This is because these companies know that they can extort more from vulnerable people—people with low credit scores, for instance—and maximize their profits as efficiently as possible.

O'Neil suggests that this shift is dangerous, because it confirms that the use of data (and the "weapons of math destruction" it powers) is stratifying society more deeply and keeping the vulnerable in dire straits, while allowing corporations and CEOs to get richer and richer. When "efficiency" is no longer for the consumer's sake but for the sake of ultra-wealthy corporations, there's a major problem. And yet, the victims of this new precedent can do little to protest or change it.

But with such an immense laboratory for analytics at their fingertips, trucking companies aren't stopping at safety. If you combine geoposition, onboard tracking technology, and cameras, truck drivers deliver a rich and constant stream of behavioral data. Trucking companies can now analyze different routes, assess fuel management, and compare results at different times of the day and night. They can even calculate ideal speeds for different road surfaces. And they use this data to figure out which patterns provide the most revenue at the lowest cost.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Page Number: 168

Explanation and Analysis

Here, O'Neil examines how trucking companies began using surveillance in their truck cabs to minimize dangerous accidents—and now use surveillance to maximize their own profits. This passage explains how the process of gathering and interpreting data for noble and useful purposes can quickly transform, as corporations can begin gathering and examining data for nefarious purposes. Of course, when the trucking companies installed cameras in their rigs, they were hoping to minimize fatalities, look out for their drivers, and make the roads safer—but because fatal accidents can cost these companies millions of dollars in insurance payouts, they were ultimately looking out for themselves.

Now, with more sophisticated ways of gathering and looking at the data these surveillance machines provide, companies can continue prioritizing profits above all else. Rather than making things easier and more efficient for their overworked drivers, they're primarily using this data to cut costs and make sure that they're bringing in as much money as possible. This is a problem, in O'Neil's estimation, because it illustrates that major companies and corporations are increasingly looking out for profits over people. Technology is being used to erase workers' humanity and agency, sacrificing fairness and transparency for efficiency and revenue.

Chapter 10: The Targeted Citizen Quotes

[Publicly held tech corporations'] profits are tightly linked to government policies. The government regulates them, or chooses not to, approves or blocks their mergers and acquisitions, and sets their tax policies (often turning a blind eye to the billions parked in offshore tax havens). This is why tech companies, like the rest of corporate America, inundate Washington with lobbyists and quietly pour hundreds of millions of dollars in contributions into the political system. Now they're gaining the wherewithal to fine-tune our political behavior—and with it the shape of American government—just by tweaking their algorithms.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 181

Explanation and Analysis

In this passage, Cathy O'Neil examines the dangerous role that tech companies play in U.S. politics. By explaining how corporations like Facebook, Google, and Amazon exert political control over the American government, O'Neil is showing how the use of "weapons of math destruction" (WMDs) directly threatens the fabric of U.S. democracy. Democratic processes should be transparent and fair—but when corporations can essentially pay off politicians in exchange for looser restrictions and tax breaks, the democratic process isn't working.

And not only are these corporations extorting American politicians—they're also using data gathered from U.S. citizens to influence how they see their government, and thus how they vote. Tech corporations can essentially weaponize data on a large scale to secure the outcomes that will benefit their business models. And when companies are using algorithms to reach people, sway their opinions, influence the government, and impact legislation, they're the ones who are in control—not publicly-elected government officials. With this, O'Neil implies that the WMDs companies use to retain their power must be dismantled to prevent them from influencing Americans.

Successful microtargeting, in part, explains why in 2015 more than 43 percent of Republicans, according to a survey, still believed the lie that President Obama is a Muslim. And 20 percent of Americans believed he was born outside the United States and, consequently, an illegitimate president. (Democrats may well spread their own disinformation in microtargeting, but nothing that has surfaced matches the scale of the anti-Obama campaigns.)

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 194

Explanation and Analysis

In this passage, Cathy O'Neil shows how microtargeting—the process of using advertising to target a person's interests based on data gathered from their internet history—is playing a major (and worrying) role in U.S. society. Microtargeting is used to show people who've already been browsing for a certain product more advertisements for related products and services. But when microtargeting is used to send certain kinds of news or information to certain kinds of people—and when that information is tied to elections and other U.S. democratic processes—it becomes dangerous.

Microtargeting is harmful because in today's world, it's used to spread misinformation. So, for instance, a political campaign could use a misleading headline or news story aimed at a swing voter (a person without a firm political affiliation) to steer that person away from the opposing candidate. This, O'Neil asserts, is how racist and dangerous misinformation about President Obama reached so many people during the 2008 and 2012 presidential; races—and why it continues to linger in people's political consciousness years later.

The combination of misinformation, invasion of privacy, and tinkering with people's points of view is warping the U.S. political system. Microtargeting is, in O'Neil's estimation, absolutely a WMD. And as such, it needs to be more carefully regulated, so that it doesn't continue to threaten the U.S.'s integrity as a democracy.

Conclusion Quotes

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 have to explicitly embed better values into our algorithms, creating Big Data models that follow our ethical lead. Sometimes that will mean putting fairness ahead of profit.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 204

Explanation and Analysis

In this passage, O'Neil argues that when it comes to the battle between humanity and technology, humanity will always be the victor. This is because we can imagine and create the futures, while machines can only "codify the past." What O'Neil is essentially saying here is that humanity can learn from our past mistakes, imagine better ways to move forward, and act with a moral conscience. But machines are unable to self-correct, unable to create, and devoid of values or ethics; they do only what they're instructed to do.

So, in the age of the Big Data economy—when companies implement computer algorithms to streamline various aspects of human life—it's up to us to restore some humanity to the processes that now govern our everyday lives. The shift toward prioritizing profit and efficiency above all else has harmed humanity and put us at the mercy of unfeeling machinery. Now, humanity must restore its role in the processes we've signed away to models and algorithms—otherwise, society will increasingly come to reflect the amoral and efficiency-focused goals of these "weapons of math destruction."

Data is not going away. […] Predictive models are, increasingly, the tools we will be relying on to run our institutions, deploy our resources, and manage our lives. But as I've tried to show throughout this book, these models are constructed not just from data but from the choices we make about which data to pay attention to—and which to leave out. Those choices are not just about logistics, profits, and efficiency. They are fundamentally moral.

If we back away from them and treat mathematical models as a neutral and inevitable force […] we abdicate our responsibility. And the result, as we've seen, is WMDs that treat us like machine parts […] and feast on inequities. We must come together to police these WMDs, to tame and disarm them.

Related Characters: Cathy O'Neil (speaker)

Related Themes:

Related Symbols:

Page Number: 218

Explanation and Analysis

Here, O'Neil approaches the conclusion of the book by reminding readers that it is a "moral" imperative to "tame and disarm" the "weapons of math destruction" (WMDs) that are increasingly exerting control over humanity.

Throughout the book, O'Neil has shown how WMDs compromise almost every sector of human life they touch—from grade-school test scores to college admissions to the hiring process to the journey of seeking credit and insurance. WMDs are now found in every part of the economy and the education system, and they're even infiltrating politics. They treat humans "like machine parts" and create staggering "inequities"—and if humanity doesn't recognize the impact that WMDs are having and do something to stop it, they'll continue taking over.

WMDs that prioritize efficiency and profit on the behalf of major corporations directly threaten people's ability to empathize with one another, to evolve as social beings, and to continue improving. By automating the processes of daily life, O'Neil argues here, humanity is acting against its own best interests—and endangering its future capacity to thrive. And yet humans have created these WMDs and decided to "pay attention" to certain kinds of data, as if algorithms and models hold the answers to how we should live (rather than oversimplifying the human experience to the point of dehumanization). It's up to humanity now, O'Neil suggests, to find a way to slowly dismantle these WMDs through regulation, oversight, and thorough vetting. Otherwise, we will be "abdicat[ing] our responsibility" to one another and to the health of our society.

SUMMARY AND ANALYSIS

The color-coded icons under each analysis entry make it easy to track where the themes occur most prominently throughout the work. Each icon corresponds to one of the themes explained in the Themes section of this LitChart.

INTRODUCTION

As a young girl, author Cathy O'Neil was a self-described "math nerd." She loved math because it was simple and neat when so much of the world was messy. While majoring in math in college and earning a PhD in algebraic number theory, O'Neil enjoyed adding to the field of mathematics, helping to expand its bounds.

By introducing herself as a person whose life has centered around math, Cathy O'Neil establishes both her credibility as a mathematician and her deep investment in making sure that mathematics is used to improve the world.

After teaching math at Barnard College for several years, O'Neil left academia for a new "laboratory"—the global economy. As a "quant" (quantitative analyst) for D.E. Shaw, a major hedge fund, O'Neil was amazed by how the operations she and her team performed each day translated into "trillions of dollars sloshing" between accounts. But in the fall of 2008, after just a year at the company, everything changed—the financial crisis brought the economy to a halt.

O'Neil hoped to use math to educate people and help them live sustainable lives. But the longer she worked with math, the clearer it became that math was an extremely powerful tool that could change the world—or derail it.

The financial collapse was made possible by people like O'Neil—mathematicians who had multiplied the "chaos and misfortune" of the crisis by misusing math. But rather than taking a step back after the crisis, people instead doubled down on new mathematical techniques—and mathematicians began studying people's desires, movements, and spending habits, calculating each human's potential as "students, workers, lovers, criminals." This, O'Neil writes, is now known as the Big Data economy.

O'Neil implies that the financial crisis should have caused people to reckon with how powerful a force math was. But instead of stopping to think about math's role in the financial crisis, corporations only barreled ahead with their use of complex mathematics in everyday operations. This situation introduces the tension between humanity and technology: without a human element to "Big Data" and technology, mathematical models threaten to dehumanize people by turning them into nothing more than statistics.

But around 2010, as Big Data saw mathematics involving itself in human affairs like never before, O'Neil began to feel troubled. People, after all, were imperfect and fallible—and the math-powered algorithms and models that were now powering the data economy had been encoded with their human creators' prejudice and biases. Yet Big Data seemed unimpeachable, even as it further deepened the wealth divide in global society. O'Neil calls these harmful models **Weapons of Math Destruction**—WMDs for short.

O'Neil suggests that mathematical algorithms and models aren't necessarily more objective than people are—after all, these algorithms are created by humans who are naturally biased. To better explain this idea, O'Neil introduces the concept of a "weapon of math destruction"—a mathematical tool that has the power to create widespread chaos in society by increasing disparities between different sexes, races, and classes of people.

One case of a **WMD** that began with an admirable goal but quickly became destructive started in 2007 in Washington, D.C. The city's new mayor, Adrian Fenty, dedicated himself to turning around underperforming schools throughout the city. Michelle Rhee, the chancellor of schools whom Fenty appointed, developed a teacher assessment tool called IMPACT that would use data to weed out low-performing teachers.

This passage introduces the idea that a WMD can have noble origins. D.C. officials just wanted to reform their school system—but O'Neil alludes to the fact that in the process, they created a situation in which a flawed mathematical model did more harm than good.

At this time, Sarah Wysocki was a fifth-grade teacher in Washington, D.C. who was beloved by her students and had consistently gotten great performance reviews from her school's principal—but she received a terrible score on her IMPACT evaluation. Because of IMPACT's algorithm, she was fired along with over 200 others area teachers. The algorithm had promised fairness—but Wysocki felt that the numbers were anything but fair.

While algorithms are often incredibly efficient, they're not always fair. In this case, an algorithm that was meant to make the D.C. school system better was inherently flawed. And because it wasn't regulated, it ended up having unfair consequences for teachers who performed well in the classroom.

The Princeton-based Mathematica Policy research had come up with the IMPACT evaluation system. They knew that measuring students' educational progress was a complex issue (and tried to make their algorithm complex, too). But they couldn't pinpoint how much of any given student's struggles in school were the result of outside factors like poverty, trouble at home, or social issues at school—and how much they reflected a bad teacher.

Algorithms and models can't solve every problem—especially when a problem is complicated and rooted in human struggles. While a model can produce data about a student's test scores, it can't know that student's individual strengths and weaknesses—nor can it evaluate how those factors are being handled in the classroom.

There were too many factors that go into the process of teaching and learning, Wysocki argued, to create an algorithm that would quantify them. Without huge collections of data points or feedback to warn statisticians when they're off-track, data models can become self-perpetuating produce results that don't necessarily tell the whole story—but that confirm whatever the statisticians set out to prove. This dynamic is highly destructive.

Wysocki essentially lost her job because an algorithm detected that her students weren't scoring high enough. This might have been efficient in terms of removing teachers whose students weren't really thriving—but O'Neil implies that it wasn't fair or even effective at solving the problem of low-performing students.

The "effectiveness" of the model that Mathematica used to weed out D.C.'s lowest-scoring teachers seemed unimpeachable. But in fact, it was an example of a **WMD** feedback loop—a situation that takes place when models "define their own reality and use it to justify their results." Other examples of this include employers who use credit scores in the hiring process, believing that responsible people have higher credit and are thus more likely to do well in a role. As a result, affluent people get jobs, while poor people get caught in a downward spiral. WMDs are dangerous because they're engineered to evaluate large groups of people, but the criteria they use to make such sweeping judgments are unknowable to everyone but their creators.

By claiming that models often create results that support their purpose, O'Neil suggests that these algorithms are effective only on their own terms. They're not fair or even particularly useful in securing accurate data and results, because they "define their own reality" by cherry-picking certain data points . So, these "weapons of math destruction" are creating widespread damage by deepening social inequality .

At the start of her last year at MacFarland Middle, Sarah Wysocki saw that many incoming fifth graders from a nearby elementary school had scored well on their tests. But when they arrived in her classroom, they struggled to read simple sentences. Later, investigations by major newspapers would reveal that there was a lot of cheating on these exams—but the students weren't the ones cheating. Their teachers, motivated by the fact that higher student test scores would reward their own performance in the eyes of evaluation algorithms, were correcting their tests for them. Wysocki would later become convinced that she herself was a victim of other teachers' desperate actions in the face of a **WMD**. The human victims of WMDs, O'Neil argues, are held to higher standards than the algorithms themselves.

WMDs can create an environment where people take desperate measures in order to avoid the algorithms' harsh judgments. In this example, scoring systems that were supposed to be evaluating student success caused fear and stress for teachers, to the point that they felt compelled to lie and cheat on behalf of their students. In this way, the standardized tests were doing the opposite of what they were created to, and they harmed both students and teachers in the process. This illustrates how destructive unregulated WMDs can be.

In 2011, O'Neil quit her job at Shaw and joined an e-commerce start-up as a data scientist. But she was disheartened to find that **WMDs** were, by now, at the heart of every industry—and they were deepening inequality everywhere. Scandalized and outraged, Shaw started taking action: she launched a blog that would expose how bad statistics and biased models were creating dangerous feedback loops. She also joined Occupy Wall Street and began speaking at the Alternative Banking Group at Columbia University, advocating for financial reform.

This passage illustrates O'Neil's investment in combating WMDs across multiple professional sectors, and in her personal life as well. O'Neil has long advocated for fairness, direct human involvement, and regulation in the use of mathematical modeling.

Yet mathematical models still control many different sectors—from advertising to schools to prisons. Models and algorithms and software only exist to grow revenue. Profits of any kind, O'Neil argues, are "serving as a stand-in […] for the truth." **WMDs** are engineered to make money or to create clout, and they ignore the people they hurt in the process—people like Sarah Wysocki. O'Neil announces her intent to take her readers on a tour of "the dark side of Big Data" and examine the injustices that WMDs cause as they control most aspects of modern life.

Big Data and the use of mathematical modeling are changing how we live. Yet people are being victimized by algorithms that prioritize profits over truth and efficiency over fairness and equality. If WMDs aren't regulated—let alone called out for what they are—society may become more divided and unjust.

CHAPTER 1: BOMB PARTS

In August of 1946, the Cleveland Indians lost the first game of a double-header. So, player-manager Lou Boudreau decided to switch up his fellow players' locations when the opposing team's greatest hitter, Ted Williams, went up to bat. Boudreau, O'Neil writes, was thinking like a data scientist—he'd analyzed Williams's hitting patterns and rearranged his own team around it. Nowadays, major league baseball is an enormously data-driven game. Baseball statisticians have spent years making mathematical models based on "every measurable relationship among every one of the sport's components."

This passage provides an example of a relatively benign approach to mathematical modeling. It shows that algorithms and data science can be used to create advantageous results without directly harming anyone. Models that measure and interpret baseball statistics aren't WMDs because they're generally both efficient and fair.

Even though baseball models now define how the game is organized, played, and bet on, they're transparent: everyone has access to the statistics that rule the game. There's a huge amount of data that's highly relevant to the outcomes that statisticians and fans are trying to predict, and it's coming in all the time (12 or 13 baseball games are played each day between April and October of each year). Many of today's **WMDs**, by contrast, are mysterious—and they often sorely lack the data for the behaviors they're interested in. So, they use proxies (stand-in data), like people's zip codes and the languages they speak, to determine things like how likely they are to pay back a loan.

The baseball models are relatively neutral because they use easily attainable data, and they stick to the numbers provided. But WMDs, through the use of proxy data and data gathered through hidden or manipulative means, are using data for their own unclear purposes rather than to measure or learn something authentically.

Even though baseball models and the model used to evaluate teachers in Washington, D.C. are incredibly different, they are both models: a model is an "abstract representation of [any] process." Human beings carry models in their heads all day—as an example, O'Neil uses the "informal model" of how she decides what to cook for her large family each night. She has data (each person's likes and dislikes), and she has new information concerning that data all the time: fluctuating grocery prices, changing tastes, and anomalies like special meals for special occasions.

O'Neil suggests that creating models and building algorithms to accomplish tasks, gather data, and streamline certain processes isn't inherently wrong. Models are useful, and it makes sense that as technology has grown more sophisticated, algorithms have become increasingly common in many sectors.

Making a model, though, requires simplification—and this means that most models have blind spots that reflect their creators' judgements and priorities. For instance, O'Neil wouldn't feed her children Pop-Tarts for every meal, even though her children love them. So, she's imposing a bias and a judgement on the model of how she feeds her family.

One of the issues with models is that they're vulnerable to human bias. This means that creators need to account for that bias rather than ignore it—doing the latter can lead to the creation of a WMD.

Models can change, too, based on their creators and purposes: O'Neil's children might build a model featuring ice cream at every meal, while a North Korean bureaucrat might optimize the model to feed a family a cost-effective bare minimum. Models reflect our personal realities—and they must constantly change or risk growing stale and irrelevant. Good models can be primitive—but primitive models can be dangerous. For example, a racist's worldview is based on a single biased and unchanging model that refuses to be affected by new "data" or experiences.

Here, O'Neil illustrates how vulnerable models are to bias, and also to faulty or outdated information and oversimplicity. Again, she's showing how much human intervention and maintenance models need in order to avoid becoming WMDs.

O'Neil turns to an example from 1997 of how racism is a brutal, unfair model. When a Black man named Duane Buck was convicted of two murders in Harris County, Texas, the jury had to decide whether he'd receive a sentence of life in prison or the death penalty. The defense attorney called a witness, psychologist Walter Quijano, to the stand—and Quijano testified that "the race factor" in the case made Buck's "future dangerousness" more likely. The jury sentenced Buck to death. The Texas attorney general would later try several cases in which Quijano's "race-based testimony" had resulted in harsher sentences—but Buck never got a new hearing, and he remains on death row.

This passage provides an example in which an unfair, outdated, inaccurate model—Quijano's racism—had devastating effects on the person to which it was applied (Duane Buck). O'Neil shares this anecdote to, once again, show how faulty, biased, and ultimately cruel models can be when they're left unchecked or fed faulty data.

Prosecutors in Harris County are three times more likely to seek the death penalty for Black people. And sentences imposed on Black men (who comprise only 13 percent of the U.S. population, yet make up 40 percent of the U.S.'s prison population) are 20 percent longer than those for white people. Courts across several states have turned to algorithms called recidivism models in hopes of erasing racism from the U.S.'s court systems. But O'Neil argues that these models can simply mask human bias.

In Harris County, there was a need for an unbiased, data-driven method of assisting judges in making tough calls. So, recidivism models that promised fairness and justice seemed like a logical solution—they'd take human bias out of the equation and thus deliver an impartial recommendation.

Recidivism models such as the LSI-R (Level of Service Inventory-Revised) are biased against poor people and racial minorities. They ask about prior involvement with the police (something that's more likely for Black and Latino youths, for instance, who are regularly targeted by police forces through programs like stop and frisk). They also examine whether one's friends or families have criminal records—again, something that's statistically more likely in low-income neighborhoods. Recidivism models, O'Neil argues, are **WMDs** because of the "toxic cycle" of damaging feedback loops and biases that structure them.

While recidivism models might claim to be race-blind and unbiased, they really aren't. Society is biased, and the model doesn't do anything to temper that. Therefore, O'Neil is suggesting that it's unjust for judges to use these models to aid in sentencing because of the implicit bias built into them. This bias can create unfair consequences—a "toxic cycle" that discriminates against and disproportionately punishes non-white people.

The baseball model and family dinner model discussed earlier, O'Neil writes, are both models that are open and transparent. But the recidivism model, she suggests, is largely invisible—and for the most part, hidden models are the rule rather than the exception. Transparency is important, and yet a hallmark sign of most **WMDs** (especially those owned by companies like Google, Amazon, and Facebook) is that they are difficult to understand "by design."

When models are used to determine outcomes as serious as prison sentences, they need to be transparently made and transparently used. But WMDs rely on being inscrutable in order to prevent people from seeing the biased or faulty ways in which they work.

Another major component of a **WMD** is its capacity to grow or scale. WMDs in human resources, health, and banking sectors are quickly becoming "tsunami forces" that define how people live their lives and whether they can access certain opportunities. Recidivism models like the LSI-R, which are presented as tools of prison reform, are perceived as being fair and efficient—but that couldn't be further from the truth.

When WMDs dictate serious decisions or when they are widespread throughout a certain part of society, they become dangerous because of how much influence they exert over humanity. O'Neil compares WMDs to "tsunami forces" to emphasize the idea that algorithms can do immense damage to people's lives.

The three elements of a **WMD**, according to O'Neil, are opacity, scale, and damage. Not all of the WMDs she'll discuss throughout the book, she writes, are entirely hidden, huge in scale, or the causes of irreversible damage. But they are all threats in those three arenas on some level. The "menace" of WMDs, she argues, is rising—and the world of finance is an example of what could happen if they spiral out of control.

O'Neil defines a WMD as an algorithm that's deceitful or unpredictable, widespread, and harmful. These qualities are why O'Neil compares harmful algorithms to weapons of mass destruction (like a nuclear bomb, for instance). Like war weapons, WMDs are a "menace" that can destabilize society, create chaos, and derail people's lives.

CHAPTER 2: SHELL SHOCKED

Quants who work at hedge funds zoom in on tiny patterns, then train algorithms to predict recurring errors and price swings, then place bets on their occurrences. The smallest of patterns can make millions for the first investor who recognizes them—and those patterns will keep raking in money until the pattern ends, or the rest of the market catches on.

Quants (quantitative analysts) who work at hedge funds essentially bet on world markets with the help of complex algorithms. Technology has streamlined the financial markets and made it easier than ever to make a profit.

While working at Shaw, O'Neil loved the "treasure hunt" component of finding market inefficiencies. At Shaw, her smarts were translating into money—lots of money. But out of the 50 quants on O'Neil's "futures group" team, she was the only woman. She was siloed from many of her other coworkers, so that if someone walked away to another firm, they wouldn't bring other quants' trade secrets with them. The work was exhausting and sometimes frightening, like when huge sums of money were needed frantically and immediately. But something deeper began to gnaw at O'Neil. The numbers she was playing with all day weren't just abstract figures—they represented people's livelihoods, retirement funds, and mortgages.

O'Neil had a moral investment in determining whether the work she and her team at Shaw were doing was legitimate and fair. They were often profiting off of uncertainty and chaos, and they were gambling with people's entire financial futures. Shaw's technology was efficient, but it wasn't always fair, which is true of all the WMDs that O'Neil discusses in the book.

In July of 2007, interbank interest rates spiked. Even though lots of people had been able to secure mortgages in the housing boom of the last several years, banks were now realizing that there was some "dangerous junk" in their portfolios. Shaw could see that many companies and world markets would suffer—but as a hedge fund, they didn't plunge into risky markets. Instead, they stood on the sidelines and bet on them. Hedge funds are less like baseball fans who cheer when their team wins, and more like gamblers who bet on movements associated with the game. So, O'Neil and her team, while nervous about what was to come, felt more or less safe. Even though the market was beginning to grow unstable, Shaw was "on top of the world."

The early days of the 2007 global financial crisis were an uncertain time, and investment firms like Shaw actually profited off of the volatility and uncertainty. By contrasting the "dangerous junk" in the banks portfolios with Shaw's position of being "on top of the world," O'Neil implies that hedge funds had an unfair advantage during this time. Even though the company was skilled and efficient at using data to make money, their processes weren't rooted in equity and justice—so their technology wasn't aimed at promoting those things either.

But even Shaw started to get nervous as the market continued to rumble. Mortgage-backed securities—previously "boring financial instruments" that could actually offset risk through quantity—became bigger liabilities. Since the 1980s, investment bankers had been buying up and packaging mortgages into securities, a kind of bond, by the thousands. The mortgages were like the "little pieces of meat of varying quality" that make up a sausage, and the securities were the spices. No one worried about them because they were essentially marked as safe. But mortgage companies had been lending money to people for homes they couldn't afford, collecting the fees, and then unloading the securities into the market. The deals that the banks were offering weren't just unsustainable—they were predatory.

Here, O'Neil explains the inner workings at banks that led to the outbreak of a devastating financial crisis beginning in 2007. The banks were creating false data through hidden and unfair methods, risking human stability (and, of course, fairness) in the process.

These subprime mortgages weren't **WMDs**—they were financial instruments, not models. But when banks turned the mortgages into securities and sold them, they relied on mathematical models—flawed ones—to do so. So, the risk model attached to the mortgage-backed securities was a WMD. None of the companies' mathematicians were updating their data and continuously balancing the risk. The numbers these companies did have had been given to them by people committing wide-scale fraud. The risk ratings were kept hidden from the public, the risk models created a feedback loop by falsely rating defective products, and the fraud was happening at an enormous scale. So, the securities had all the components of a lethal WMD.

WMDs can spin out of control when there's no transparency, accountability, or regulation. It's dangerous for a model to operate without being regularly updated—especially a model like the risk models that were keeping the American (and global) economies afloat. Because the risk models seemed to be efficient in terms of making a profit and keeping things calm on the surface, they weren't regulated or examined, which allowed them to become WMDs.

The algorithms that had created the market and analyzed the risk in the securities turned out to be useless. Disaster hit the economy, and the human suffering it created was finally on display. In the financial sector, everyone—including the quants at O'Neil's firm—began to wonder what would happen next. But by 2009, it was clear that the industry hadn't really learned anything or changed—there were just a few more hoops to jump through.

In the throes of the financial crisis, only humans could sift through the mortgages and assess the true values of the loans. Technology was supposed to be able to run itself—but in the end, a human hand was needed to mitigate the disaster that had struck. Yet as the crisis dissipated, no one seemed motivated to investigate how to prevent something similar from happening again.

O'Neil had become disillusioned with the world of finance; people were wielding formulas recklessly and inappropriately. O'Neil left Shaw in 2009, planning to work on fixing **WMDs** from the inside out by joining a group that provided risk analysis for banks. The longer she worked as a risk analyst, though, the more she got the sense that she and her colleagues were seen as "party poopers" or threats—even given the cataclysmic crash that the country had just been through.

Here, O'Neil shows that even after switching industries, she still wasn't seeing anyone express the desire to take accountability for how algorithms were destabilizing the world. Furthermore, no one was taking the initiative to investigate how to stop them from growing further out of control.

In 2011, O'Neil switched roles yet again. She joined a web start-up as a data scientist, where she built models to anticipate the behavior of the users who visited travel websites and to try to distinguish casual window shoppers from motivated buyers. As she adjusted to her new job, she found lots of parallels between finance and Big Data—the biggest of which was that in both fields, money and self-worth were inextricably interwoven. People believed that they were successful in finance because they deserved to be, rather than because they were simply lucky. This itself was an example of a feedback loop. "Money," O'Neil writes, "vindicates all doubts."

This passage underscores how tech companies and financial institutions alike prioritize efficiency—which often translates to the most profits for the least amount of hassle and work—over fairness, justice, or empathy. This focus on effectiveness over everything leads to the proliferation of WMDs throughout many industries.

O'Neil continued to grow disillusioned by how her new industry sought to replace people with data trails and turn them into more "effective" voters, workers, and consumers. She could see a new kind of dystopia growing around her, and inequality rising nonstop as data controlled and manipulated more and more people. Eventually, she quit her job to devote more time to investigating how algorithms were destroying lives.

O'Neil resisted society's increasing emphasis on efficiency and effectiveness over fairness and justice. Humans were being turned into data points, and too many different parts of their lives were being altered and streamlined. WMDs, in O'Neil's estimation, had begun to take over the world.

CHAPTER 3: ARMS RACE

To explain one of the core components of a **WMD**—scale—O'Neil invites her readers to imagine that the trendy "caveman diet" became the national standard, and all 330 million Americans were forced to follow its dictates. The restrictive diet, which favors meats, fish, fruits, vegetables, nuts, seeds, and cheeses, would have a huge effect on the economy—if it became a national standard, it would create a distorted economic climate. This is precisely what has happened to higher education.

This passage suggests that standardizing certain protocols or ways of living can be harmful. Just as parts of the economy would collapse if certain major exports were suddenly undesirable, the higher education sector has created chaos by enforcing certain standards.

In 1983, the *U.S. News & World Report*, a struggling magazine, decided to evaluate and rank 1,800 colleges and universities across the U.S. to bring in readership. They based their rankings off opinion surveys sent to university presents—but after the first rankings were released, complaints started pouring in, and the editors at *U.S. News* tried to figure out how they could statistically measure the vague concept of "educational excellence." So, they decided to look at SAT scores, acceptance rates, student-teacher ratios, alumni donations, and more, building an algorithm to create the first data-driven ranking in 1988.

Just as it was unfair for the IMPACT algorithm to try to sort D.C.'s good teachers from its bad ones, the U.S. News *standards were failing to measure nuance. Trying to gather data that would make their rankings fairer and more transparent was a step in the right direction—but they were relying on algorithms and models rather than human subjectivity.*

The rankings, though, quickly became mired in a feedback loop—low-ranked schools received fewer applications from top-tier students. Thus, their percentage of high-achieving students dropped, as did their revenue from admissions and alumni donations. And the *U.S. News* list kept growing in reputation and scope—soon, it was a "bona fide **WMD**."

Whichever schools U.S. News *said were the best quickly became the most attractive—while those they ranked lowest suffered. This is an example of a destructive feedback loop, in that a low ranking (which could be based on biased or otherwise faulty data) leads to a lower reputation—which, in turn, reinforces the low ranking. This cycle is a hallmark of a WMD.*

Many schools began looking for ways—both legal and extralegal—to bump themselves up in the rankings. But the *U.S. News* model was constructed from proxies—the most relevant data about students' day-to-day experiences at their colleges wasn't accessible. So, it was easy for colleges to improve in the arbitrary areas the ranking measured. But as universities hustled to improve faster than their competitors, a kind of "arms race" began. Texas Christian University poured money into its student center and its football team to attract more students. Their strategy worked—by 2013, it was the second-most selective university in Texas. It climbed 37 places in the ranking in just seven years.

This passage shows that the rankings did create some good—they inspired schools that weren't ranking well to make major improvements to their organizations and infrastructures. But still, schools were being ranked based on stand-in data rather than the most up-to-date information, so schools were being judged based on potentially faulty algorithms.

The rankings were creating a kind of rat race in U.S. academia. As more and more schools adopted algorithmic approaches to things like admissions, safety schools started rejecting the top-tier candidates whom they knew would likely reject offers for better universities. So, safety schools were no longer "safe." O'Neil argues that it's not just students who are suffering in this new climate—it's the schools, too, who are losing out.

As the U.S. News *algorithm became the standard in American higher education, American schools began adopting algorithms of their own that would keep them competitive. In this way, they were sacrificing fairness and transparency for efficiency—essentially, they were being governed by WMDs.*

By leaving tuition, fees, and student financing out of their initial model, O'Neil argues, *U.S. News* ultimately did its readers an enormous disservice. By ranking expensive, prestigious universities highest, they implied that money was the cost of excellence. And indeed, between 1985 and 2013, the cost of higher education rose by over 500 percent—four times the rate of inflation. And while the student loan crisis isn't the fault of *U.S. News* alone, the idea that a degree from a highly ranked school is a path to power and success is partially their doing.

The U.S. News *rankings are an example of a WMD in that they essentially drove up the cost of American education through their algorithm. They created impossible standards, using efficient but unfair algorithms, and the entire U.S. economy—and countless students—suffered. The* U.S. News *rankings are still creating inequality and a deeper class divide in the U.S. by making higher education less accessible.*

Colleges manage student populations "like an investment portfolio," according to O'Neil, by assessing which students are assets (those who pay full tuition or have families who donate money) and which are liabilities (student athletes who receive lots of scholarship money to attend). Education consulting firms have sprung up to do the analytical work of forecasting and ranking enrollment prospects by a number of categories and characteristics—often using the *U.S. News* metrics as a model. All over the U.S., rankings—and efforts to game them—continue to grow and spread.

This passage continues to show how data and algorithms have all but taken over higher education, essentially erasing humanity from the equation. Entirely new businesses have sprung up to take advantage of this increasingly divided climate in the higher education sector, creating harmful new algorithms of their own. Higher education, particularly in the U.S., is now mired in what O'Neil implies is an unfair rat race.

To try to level the playing field, Saudi universities paid successful academics exorbitant fees to come on as adjunct faculty. And students in rural areas of China protested in defense of their right to cheat on standardized tests (to even things out between them and their urban counterparts). It was becoming clear that the rankings had created a system that many people felt there was no way to win fairly. Expensive application bootcamps and admissions coaching sessions now dominate education by turning students' test scores and GPAs into statistics. And colleges and universities create admissions models that are hidden, large-scale, and trapped in feedback loops—they are **WMDs**. The contemporary education system favors the privileged—those who have the means to play by the algorithms' rules.

Students, educators, and institutions are all trying desperately to play catch-up and compete with the algorithms that now efficiently but unfairly decide people's entire educational futures. Because the algorithms found in higher education are widespread, opaque, and damaging, they're creating vast social destruction that might never be repaired.

During President Barack Obama's second term, he suggested a new college rankings model that was tied to affordability, diversity, and postgrad job placements—but any ranking system, no matter its priorities, can still be gamed. For instance, a law school graduate with hundreds of thousands of dollars in loans working as a barista is considered employed. Another example is that schools can lower costs by replacing retired tenured professors with overworked adjuncts who cost less to hire.

Even though there have been initiatives to try to combat WMDs in higher education, these algorithms seemingly can't be beat. The drive toward efficiency and profits over empathy and humanity is making it more difficult for people to attain quality higher education, adequate financial assistance, and fair working conditions.

The government failed to rewire college rankings. Instead, the Education Department released lots of data online, hoping that students would use the data to determine what matters to them individually about their collegiate experiences. The website allows people to create models for themselves that are transparent, user-controlled, and personal. O'Neil suggests that the Education Department's new site is "the opposite of a **WMD**."

This passage shows a possible solution to WMDs: putting agency back in the hands of people and allowing them to model their own futures. O'Neil suggests that making algorithms more transparent and participatory is the only way to stop WMDs from gaining more and more control over society.

CHAPTER 4: PROPAGANDA MACHINE

While working as a data scientist at an advertising start-up, O'Neil and her team hosted a visit from a venture capitalist who gave a speech describing the "brilliant future" of targeted online advertisements. Consumers would contribute data through their online behavior, and advertisers would target them with "valuable information" that would help them shop and live better. No longer, he joked, would web users find themselves assaulted by ads from the University of Phoenix. Though O'Neil's coworkers laughed at the joke, she was perturbed—the internet was already preying on lower-income people with "the bait of upward mobility."

While online advertising targeted at individual users is now common, O'Neil suggests that this doesn't mean it isn't predatory—and it's especially dangerous for low-income people who are already at a disadvantage. The University of Phoenix is a for-profit college, which means that its goal is to make money rather than invest back into resources for students. So, baiting low-income people with the promise of "upward mobility" (increased class status) if they attend a college like this could put them at an even greater disadvantage, because they may not end up receiving the quality of education that they're looking for. In this way, WMDs like targeted advertisements are deepening social inequality slowly but steadily—even as their creators joke about them.

O'Neil alleges that the internet isn't the equalizing, democratizing force it promised to be. Instead, as Big Data and tech companies have learned more about individual users, they've created rankings and categorizations of people that target their vulnerabilities, feed them predatory ads, and exploit their finances. Online, for-profit universities that charge exorbitant fees advertise and direct their recruiters to people who are on government assistance, who've been recently incarcerated, or who are otherwise financially or socially vulnerable. "Vulnerability is worth gold" to advertisers and the makers of **WMDs**—and recruiters at places like ITT Technical Institute are told to "Find Out Where Their Pain Is" when they're engaged in recruiting.

Online advertisers know that they're deepening the gulf between the rich and the poor. But because their algorithms are aimed at maximizing exposure and profit, they don't take fairness into account. In fact, they exploit existing social problems like economic inequality through their algorithms, trapping desperate people in dangerous feedback loops that actually make it harder for them to advance and improve their lives.

Big tech companies like Google and Facebook allow these for-profit universities to segment their target populations and advertise to them directly. The ad campaign then runs competing ads against one another to see which bring in the most users. This method is based on A/B testing. An example of A/B testing is credit card offer mailings. Credit card companies send out massive mailings to hundreds of millions of people, so that even if only a fraction of a percentage of people respond to the ad, the company is still raking in many customers. Moving these campaigns online, though, allows advertisers to track feedback much more closely and zero in on their most effective messaging.

By casting wide nets, tech giants and advertisers make sure that virtually no one is spared from their algorithms' influence. Again, A/B testing is an example of a WMD—and it's having devastating consequences on society by roping vulnerable people into predatory schemes driven by unregulated algorithms and data.

Now, data-crunching machines can learn as they operate. This phenomenon, called machine learning, has been studied since the 1960s, when language scientists began teaching computers to read. As the internet has grown and expanded, users have given machines "quadrillions of words" about the way humans live—data machines now have one of the biggest sample sets of raw learning material ever. As advertising programs learn more and more words, they can probe users for deeper patterns.

Computers can "learn" to a certain degree—but they're mainly learning to target and exploit people. This illustrates how pivoting the economy toward data, algorithms, and maximizing profits at any cost is hurting rather than helping society.

Another kind of predatory online targeting called "lead generation" uses falsified information and misleading phrases and imagery to come up with lists of prospects that can then be sold to places like for-profit universities, who then use the user information to recruit people. Fake job postings and misleading promises of routes to food stamps or Medicaid coverage target vulnerable people, gather their information, and send it to recruiters who then make follow-up calls and emails. For-profit colleges also use financial aid questionnaires on websites like the College Board to advertise themselves to those most in need of financial assistance to attend school. Then, they arrange free online resume-writing workshops in order to harvest students' data and essentially "stalk" them with cold calls, emails, and more ads.

Here, O'Neil shows how sophisticated but predatory advertisements exist only to gather data that can then be used to power even more predatory algorithms. By specifically targeting low-income people who are eager to better their social standing, these advertisements actually deepen social inequality by influencing already struggling people to spend money on degrees that may not benefit them. So, these algorithms are essentially taking money from vulnerable people in order to maximize the company's profits.

WMDs are damaging to peoples' lives. But in the case of for-profit colleges and online advertising, the damage doesn't begin until students targeted by advertising take out loans to pay their tuition and fees. Even though the Higher Education Act of 1965 states that colleges cannot get more than 90 percent of their funding from federal aid, for-profit colleges exploit this stipulation by helping poor and vulnerable students line up massive loans. By creating highly refined WMDs that target the poorest 40 percent of the U.S. population, these colleges take advantage of desperate people, take their money, and hand out essentially worthless degrees that don't always enable them to get better jobs. In doing so, they create a damaging, destructive feedback loop.

Again, these complicated WMDs are further dividing an already stratified American society. They're promising to help people improve their lives—but really, they're only keeping poor people poor while making rich people (and corporations) even richer. This kind of feedback loop isn't sustainable—if these WMDs are left to operate on their own without any regulation, they will do irreparable damage to individual lives and society more broadly.

Loan companies, too, operate **WMDs** in order to target and draw in customers. Some of these companies are legitimate—but many charge exorbitant interest rates, sell their customers' data, or even hack their bank accounts. As a result, lawmakers are pushing for legislation that will govern the growing market for personal data. Yet many "effective and nefarious" WMDs, O'Neil writes, have no problem creating workarounds that will allow them to study our behavior online and off, carefully and relentlessly.

WMDs have been growing unchecked and unregulated for so long that it's now extremely difficult for even government policy to reel them in. The algorithms are very sophisticated—and there's always a new method of gathering and using data that works around attempts to protect society's most vulnerable.

CHAPTER 5: CIVILIAN CASUALTIES

In 2013, after the recession forced the city of Reading, Pennsylvania to make cuts to is police force (despite persistent crime), police chief William Heim invested in crime prediction software. The software was called PredPol (short for "predictive policing"), and it was made by a Big Data start-up based in California. The software promised to use historical crime data to show, hour by hour, where and when cries were likely to occur. By patrolling these hotspots, police could potentially cut down on crime—and a year later, burglaries in vulnerable areas were down by over 20 percent.

This passage continues to illustrate how technological advancements promise to make many parts of contemporary life more efficient. But thus far in the book, O'Neil has given several other examples of new technologies that have removed human influence from their processes and prioritized optimization over fairness and thorough vetting. It's likely that this will be the case for PredPol as well.

Predictive policing software operates the same as a lot of baseball statistics modeling—and because it targets geography, it's supposedly free of the racism and biases that are embedded in the recidivism models that the court system uses. And yet by allowing police to hone in on "nuisance" crimes—vagrancy, panhandling, and small-scale drug use—police are more likely to target vulnerable and impoverished areas. When officers over-police an area, the policing creates new data, which justifies more policing—and so this software can create dangerous feedback loops.

Since predictive policing software drive officers to certain geographic hotspots, they may claim that their policing isn't motivated by profiling. But at the same time, policing an area heavily means that it's more likely to show up as a place where crimes are being stopped—so the software will automatically redirect police to these neighborhoods (which are often low-income and/or predominately non-white).

Human programs like stop and frisk (a program in which NYPD offers were given the go-ahead to stop, search, and frisk anyone who seemed suspicious anywhere at any time) have been shown to create even more friction and danger in vulnerable communities. Mathematical models now dominate law enforcement. And because of theories that link nonviolent crimes to a proliferation of violent crimes, police chiefs tend to believe that even nuisance data is useful in creating "better data" that could be used to focus more heavily on violent crimes.

Stop and frisk was a program that encouraged racism and other forms of discrimination, since officers were encouraged to stop and search people based on their own subjective judgments and biases. Predictive policing promised to solve this problem by taking human prejudice out of the equation, ideally making policing unbiased and fair.

Predictive policing software, however, doesn't have the capacity to predict white-collar crime—even though the crimes carried out by the rich in the early 2000s arguably created some of the most widespread devastation in the U.S.'s recent history. Police forces across the country are using predictive policing data to reinforce zero-tolerance policies for violent crimes and nuisance crimes alike. They focus almost exclusively on the poor, so it's clear that they have a choice in where they direct their attentions. PredPol is, according to O'Neil, essentially a "do-it-yourself **WMD**." Its inner workings are hidden from the public, it creates dangerous feedback loops, and it's growing in scale.

Certain crimes, like tax evasion or money laundering, aren't the kinds of crimes that predictive policing software is looking out for. So, the software—like the human police who use it—is biased toward nuisance crimes committed in low-income or minority neighborhoods. Thus, predictive policing software is, at its core, biased against people of color and working-class people—even though it claims to be working against racism and classism.

While attending a data "hackathon" in New York in the spring of 2011, O'Neil and the New York Civil Liberties Union worked to break out important data on the NYPD's controversial and harmful stop and frisk program. Data was driving police to stop, search, and frisk more and more people—mostly Black and Latino youths, only 0.1 percent of whom were actually linked in any way to a violent crime. Stop and frisk itself, O'Neil writes, isn't a **WMD**—but it uses calculations to excuse thousands of invasive stop and frisk instances in vulnerable neighborhoods. Stop and frisk, while run by humans, did create terrible feedback loops, punishing Black and Latino men disproportionately for petty crimes and misdemeanors (like drinking in public) that white people were rarely punished for.

Here, O'Neil recalls an instance when she and her colleagues worked together to try to expose the data behind stop and frisk—in a sense, they were restoring a human hand to an automated algorithm. Reviewing how and why data encourages police to stop and frisk people in certain areas is essential in order to make sure that data isn't contributing to discrimination (like racial profiling).

Non-white people who live in poorer neighborhoods with minimal access to good schools and job opportunities are more likely to be highly policed. So, **WMDs** like predictive policing and recidivism models used for sentencing guidelines are inherently racially biased and logically flawed. Even though these models claim to have lots of data about how likely a given person from a certain neighborhood is to commit more crimes upon release from prison, they don't take into account the human factor. Thus, these WMDs only justify the systems that already exist—they don't actually gather the data needed to question or improve them.

O'Neil suggests that data-driven programs like stop and frisk and PredPol can worsen the police's unequal treatment of white and Black or minority people. While these WMDs advertise themselves as fair, they're maligned by a bias against racial minorities and low-income people.

O'Neil suggests that data scientists for the justice system should actually learn what goes on inside prisons, and how those experiences affect prisoners' behavior. Solitary confinement, rape, and malnutrition are all huge problems in the prison system. But a serious data scientist, O'Neil suggest, would look at how things like more sunlight, more sports, better food, and literacy or educational programs might impact recidivism rates. Yet private prisons make up a $5 billion industry that thrives only when its institutions are at capacity. So, rather than analyzing what goes on inside prisons, these companies purposefully work to make prisons mysterious spaces.

Rather than just looking at what factors make someone likely to commit a crime, O'Neil's suggests that people should start looking at what actually makes life fairer, easier, and more equitable for disadvantaged people. Without transparency and a sense of humanity, any data-driven approach to crime and punishment will only contribute to greater inequality.

Stop and frisk, O'Neil suggests, will soon be a thing of the past. Facial recognition software is evolving every day, and soon, data-driven approaches to spotting potential lawbreakers will breed even more destructive **WMDs**. Already, police departments around the country are employing technology experts to develop WMDs that attempt to determine which people are most likely to commit crimes.

Technology is only going to get more sophisticated as time goes on—yet O'Neil implies that a serious reckoning with the social consequences of using biased data might never come to pass. So, as technology gets more advanced, it threatens to deepen social divides even more greatly.

In 2013, a 22-year-old Chicago man who lived in a high-crime, low-income neighborhood received a knock on his door from the Chicago PD. They told him that the force had their eyes on him, since he was associated with people who'd been caught up in the criminal justice system. Rather than trying to get to know the people who lived in neighborhoods where crime was an issue, the police were deepening divisions by essentially spotlighting innocent people.

This anecdote illustrates how data-driven surveillance can encourage discrimination. Innocent people can find themselves in the crosshairs of the police simply because an aggregation of data indicates that they're similar to people who have committed crimes in the past. WMDs are directly creating painful incidents like this one.

It's simpler, O'Neil asserts, to gather data and build models that assume people are all the same than to pioneer programs that help make the justice system fairer (though perhaps less efficient). Poor people and people of color all over the country, she says, are being caught in "digital dragnets"—and in the meantime, the affluent and white people who don't trigger the algorithm get to live in blissful ignorance.

The algorithms that try to streamline criminal justice are doing serious and perhaps irreparable damage to U.S. society by dividing people along the lines of racial and class.

CHAPTER 6: INELIGIBLE TO SERVE

After a year and a half away from school to seek treatment for bipolar disorder, college student Kyle Behm had a friend recommend him for a job at a grocery store. But Kyle didn't get an interview—and his friend later explained that Kyle had been "red-lighted" by a personality test he took as part of his application. Kyle applied to other jobs, and he was rejected from every single one. Kyle's father, Roland, was an attorney. When he learned about what was going on, he sent notices to seven companies announcing his intent to file a class-action lawsuit, alleging that using the exam to weed out job applicants was unlawful.

Job-related personality tests are another form of WMDs that gather data about people and use that data in ways that are secretive or difficult to understand. They prioritize efficiency over fairness as they sort people into groups based on raw data alone. The job application process may be more streamlined in the present day compared to the past, but people aren't necessarily being assessed fairly.

WMDs aren't just corrupting the college admissions process or the criminal justice system—they're hurting jobseekers, too. While looking for a job, people used to turn to their networks of friends and acquaintances for connections—so jobseekers who weren't well-connected struggled to find work. Companies like Kronos that brought science into human resources in the 1970s promised to make the process fairer and eliminate some of "the guesswork" in hiring. But now that the hiring business brings in $500 million annually and employs tests and algorithms to weed out applicants, the job market is more unfair than ever before.

Personality tests are essentially used to exclude as many people from the hiring process as possible in order to make that process simpler and more streamlined. However, these tests sacrifice fairness for the sake of efficiency, because they ask leading and invasive questions that don't necessarily determine what kind of worker or colleague a person would actually be. People's applications—and indeed their lives—suffer as a result.

The problem with the use of these personality tests in the hiring process, O'Neil states, is that no one knows what the tests are looking for—the process is completely mysterious, the models are rarely updated or investigated to see how they're excluding people, and they are increasingly common.

Because the methods and reasoning behind these personality tests are hidden, large-scale, and responsible for harmful feedback loops, they qualify as WMDs in O'Neil's estimation.

Unsurprisingly, racial and ethnic minorities are most vulnerable to these application algorithms' fallibilities. In the early 2000s, researchers from the University of Chicago and MIT send out 5,000 fake resumes for job openings at respected news outlets for a number of different roles. Each resume was modeled for race—half featured stereotypically "white" names, like Emily and Brendan, and others featured stereotypically "Black" names, like Jamaal and Lakisha. The "white" resumes got 50 percent more callbacks than the "Black" ones—and even "Black" resumes featuring strong qualifications were rejected outright. Even though it was automated, the hiring market was still "poisoned by prejudice."

This anecdote illustrates how racial bias is often encoded in the job-hunting process. It furthers O'Neil's argument that humanity needs to reassert its presence in some of these algorithms—many of which have gotten out of control over time. Encoded biases have begun to dictate outcomes like who doesn't and doesn't get hired at certain jobs, which results in racial discrimination.

Human resources departments rely on these automated processes to help sift through huge numbers of resumes. Because these algorithms prioritize certain buzzwords and skills, they're changing the way jobseekers write their resumes and cover letters. And those who have the money, time, and resources to prepare their materials based on insider information as to what the algorithms are looking for are those who wind up winning roles. At the same time, those from poor or disadvantaged communities often lose out.

The algorithms that now define the job-hunting process were meant to remedy human biases that can contaminate the process and create inequality. Yet now, these WMDs continue to treat people unequally—and they're doing so on a larger scale than ever before.

In the 1970s, the admissions office at St. George's Hospital Medical School in London needed a way to handle the massive number of applications they received each year. Administrators created a model that they believed would boost efficiency in culling applications while remaining fair and objective—but the very inputs that the humans taught the computer were biased and racist. The model they created excluded applicants with foreign names and rated female applicants lower in the system.

This passage illustrates how human bias can infiltrate even seemingly fair or objective algorithms. In this instance, the admissions model perpetuated racism and sexism—the opposite of what the admissions office set out to do.

In 1988, the British government found the school guilty of discrimination. Rather than axing female and immigrant applicants for whom childcare and language barriers might have been struggles, O'Neil suggests that St. George could have helped these worthy candidates and provided them with resources to make their careers easier and better. **WMDs** could help lots of people, but instead, they often serve unfair objectives.

Here, O'Neil illustrates how easy it would be for WMDs to work in the opposite direction—by singling out women and immigrants not to discriminate against them, but to elevate them and help them succeed. WMDs don't have to be divisive and destructive. However, the humans that create and use these algorithms are often confirming or upholding their own biases, whether they intend to or not.

The objective of the **WMDs** created to filter out job candidates is almost always to reduce the risk of bad hires and cut down on administrative costs—in other words, they're designed only to save money. These WMDs also try to filter out which candidates are most likely to stay at a job for a long time, preventing turnover (or "churn") in the workplace. Those whose resumes reflect short stints at previous jobs or suggest that they're a more "creative" or free-spirited type will score lower in some of these algorithms.

These WMDs are meant to efficiently and quickly find the most potentially reliable job candidates. But in the process of using data to sort people from one another, they punish applicants for their uniqueness and automatically reject potentially competent and talented prospective employees.

These **WMDs** took another dangerous thing into account: commute time. By removing access from applicants who live farther away from their jobs, these WMDs are directly contributing to feedback loops that keep poverty and immobility alive. Eventually, some companies did remove this metric from their models.

If a workplace is located in a high-income area, it's usually the case that low-income employees will have to commute there. So, by ignoring candidates who lived far from their place of work, these algorithms were essentially telling low-income candidates that they weren't worthy of the job. This perpetuates poverty, as it prevents those seeking upward mobility (higher income and class status) from reaching new opportunities.

Employers also tend to want to see whether someone will be a "team player" or not. Social networking sites like LinkedIn provide a glimpse into people's social and work relationships. But Gild, a San Francisco-based start-up, sorts through millions of job sites to collect and analyze "social data," attempting to quantify and qualify workers' "social capital." By tracking what sites people visit and what kinds of social media they engage with in their off-hours, Gild seeks to learn more about what kind of person an applicant is.

Even though Gild is trying to paint more holistic portraits of job applicants, their approach is problematic because busy workers who have families or offline pursuits and hobbies may not register as being sufficiently active and social online. Again, Gild's algorithm is discriminatory, and it prioritizes efficiency over fairness.

Hiring models, O'Neil asserts, are prone to confirmation biases rooted in "pseudoscientific nonsense." In this way, modern-day Big Data is a lot like phrenology—the racist and long-debunked study of whether irregularities of the human skull signaled things about personality and destiny. Like phrenology, a lot of the tenets that motivate Big Data are "untested assumptions."

Even though the Big Data economy claims to offer solutions that will make life simpler and fairer, its methods aren't vetted or regulated in any way. Thus, they have the potential to become harmful, widespread, and destructive—just like phrenology, a racist methodology that lumped certain groups of people together based on faulty logic.

CHAPTER 7: SWEATING BULLETS

American workers have recently coined a new idea: "clopening." A combination of the words "closing" and "opening," clopening refers to when an employee works late closing one night and comes in early, sometimes just a few hours later, to open up shop the next morning. While having the same employee or employees close the store one night and open it the next morning makes logistical sense for employers, it can create stress and sleep-deprivation for workers. And because retail and food service schedules often arrive on short notice, some employees might only find out a day or two in advance that they have one or more clopenings coming up.

Here, O'Neil describes a "clopening"—a scheduling quirk that is arguably unfair for employees yet efficient for employers. Though it's not yet clear how this example directly relates to WMDs, it illustrates the same prioritization of efficiency over fairness that's common in the Big Data economy.

Irregular schedules and clopenings are both products of the Big Data economy. **WMDs** that treat workers "like cogs in a machine" create these trends and entrench them into the workplace. Scheduling used to be driven by human observation: if an employee at a family-run store noticed that there were no customers on Tuesday morning but a huge rush on Saturday afternoons, the shop might close Tuesdays and hire additional workers for the Saturday shift. But now, businesses use software to analyze customer traffic, determine exactly how many employees need to be in-store and when, and keep staffing (and spending) at a bare minimum. Gone are the days of student workers studying during downtime on the job—now, every moment of every workday is analyzed and scheduled for maximum efficiency.

By removing the human perspective from scheduling, technology has perhaps made things more cost-efficient for employers. But in the process, it's made things unfair for workers. When every aspect of a business is optimized for maximum profit, the people who keep the business running become less important than the pursuit of efficiency.

U.S. government data shows that over two-thirds of food service workers and over half of retail workers find out about scheduling changes with less than a week's notice. When *The New York Times* ran a 2014 article about a Starbucks worker named Jannette Navarro—a single mother struggling to work her way through college—Starbucks promised to change its scheduling practices by eliminating clopenings. But a year later, Starbucks hadn't made good on their word. Minimal staffing was baked into their company culture and operations. Inefficiency is a huge liability at chains like Starbucks, and individual managers could be punished for a downturn in revenue related to inefficient scheduling.

Once businesses start prioritizing efficiency over their employees' well-being, it's hard to stop. Some of these businesses, especially large ones like Starbucks, have models that only make sense profit-wise when workers are exploited so that the maximum profits can be achieved. In this way, O'Neil implies that WMDs are changing the face of global labor practices for the worse, creating irreparable damage.

Modern-day scheduling technology is rooted in the discipline of applied mathematics called "operations research"—OR for short. Mathematicians used to use OR to help farmers plan crop plantings. During World War II, OR was used to help the U.S. and British militaries optimize their resources. After the war, OR was used in manufacturing and supply chain logistics, and now it underpins huge companies like Amazon, FedEx, and UPS. But these models exploit workers, bending their lives to unfair schedules. Optimization programs are everywhere now, and they've contributed to the creation of what O'Neil calls a "captive workforce."

Even though scheduling software was originally created to help make workers' lives better and easier, it has become a WMD. It's widespread, it keeps workers in the dark, and it creates a feedback loop in which workers can't better their lives because they're beholden to long, irregular hours. Again, a model created to prioritize efficiency over every other metric has spun out of control, erasing consideration for humanity from its processes.

These over-optimized schedules create anxiety, sleeplessness, and stress that keep workers down, no matter how often they switch jobs in search of a better system. And when workers' schedules are in chaos, their children grow up without routines. In this way, **WMDs** are seriously affecting people that they shouldn't even touch. Efficiency outweighs goodness and justice—and this, O'Neil asserts, is the very nature of capitalism.

This passage shows how WMDs' effects trickle down from generation to generation, derailing innocent people's lives and creating dangerous feedback loops in the process. By prioritizing efficiency, these models are creating chaos and pain for workers—and for their families, as well.

In 2008, a company called Cataphora created a software system that used information gathered from employees' corporate emails and messaging systems to determine what kind of workers they were. People who sent emails that others copied and pasted a lot, for instance, could be seen as ideas generators; other workers were the "neurons" that connected people and only transmitted information.

Again, O'Neil uses this anecdote to show how relentless attempts to optimize profits, single out efficient and productive workers, and cull any employees who aren't massively efficient all the time are creating destruction and unfairness across the work force.

As this type of analysis became more popular in workplaces around the country, it started to have terrible consequences. Call-center employees were monitored so that their tones of voice and speech patterns could be analyzed for efficiency. And when the 2007 financial crisis hit, less-useful employees across the workforce were laid off based on the software's determination of their work styles were useful enough. This was the result of "digital phrenology."

The programs that facilitate invasions of privacy, like monitoring calls and emails, are WMDs. In particular, the example of these programs being used to lay off employees during the financial crisis shows how damaging these data-driven programs can be, since companies will do anything to stay afloat in tough times.

People who get fired because of these kinds of algorithmic metrics don't always deserve to lose their jobs—but because they do, the algorithm is reassured that it's working, and its criteria become even more entrenched. Now, tech workers and creative types are often beholden to the same crude analysis and efficiency measures that dictate the lives of overworked retail and service workers.

Software designed to analyze employee efficiency is harming the modern workplace. The feedback loops that these WMDs create are unvetted and unreliable—yet they're being used to excuse permanent changes in the ways people work and the standards to which they're held.

In 1983, the Reagan administration warned of a "rising tide of mediocrity" in American schools—SAT scores seemed to be plummeting. The administration's report suggested that underperforming teachers were the cause of the drop, and that they needed to be weeded out. This alarm was the root of programs that would derail the lives of people like Sarah Wysocki. Teachers are workers too, and they are extremely vulnerable to **WMDs**.

Here, O'Neil switches topics to explain how today's WMDs are rooted in the past. In this particular example, she draws a connection between educational practices in the 1980s and the present-day obsession with efficiency and optimization at the cost of fairness. In doing so, O'Neil implies that programs like IMPACT (which cost Wysocki her job) may seem modern and sophisticated, but they're actually based on arguably outdated concepts like the Reagan administration's "rising tide of mediocrity."

When Tim Clifford, a middle school English teacher in New York City, scored abysmally on a model's evaluation of his performance, he was devastated—but soon, he found out that many of the educators he'd worked with for years were scoring dangerously low, as well. He had tenure, so his job was spared—and the next year, his score shot from a miserable 6 to a brilliant 96. Clifford now knew for sure that the scores were arbitrary and "bogus"—yet they were threatening to ruin teachers' lives.

Clifford's experience shows how faulty and imprecise algorithms are toying with teachers' livelihoods. These models claim to create greater efficiency and better outcomes for students, but their creators fail to ensure that the models are functioning fairly, let alone accurately.

Researchers and analysists following up on the Reagan-era panic in the later 1980s found that the initial outcry was unjustified. Earlier analysts had overlooked the fact that lots of factors (e.g., more students taking the test, and universities opening their doors to more diverse student bodies) were having an impact on changing scores. When these new researchers divided the scores up into subgroups based on economic status, they found that they weren't dropping that sharply at all. The phenomenon behind this misinterpretation is called Simpson's paradox, in which a body of data displays one trend when taken as a whole but shows the opposite trend when broken up into groups.

Again, emphasizing that the Reagan administration's panic over a "rising tide of mediocrity" was uncalled for, O'Neil suggests that present-day evaluations based on this panic are inherently flawed. In addition, researchers' failure to thoroughly dissect and analyze the data behind teachers' performance scores means that teachers were unfairly punished.

Botched statistics led to Tim Clifford's struggles, too. His scores were random. In the effort to make sure that teachers were being measured based on how much they were helping their students to improve, the value-added model being used to gauge success tried to predict what a student's score would be and reward or warn teachers based on the gap between the expectation and the reality. But because teachers' classes change every year—and because a class of 25 or 30 students isn't a big enough data set, the result is a model that is essentially "noise."

Value-added models are, like many other models and algorithms, aimed at efficiently quantifying a metric that's difficult to measure: success and growth. But ultimately, these models are failures, because they simply can't work the way they're designed to. Yet with no oversight or regulation, these models are used to determine which teachers will keep their jobs and which won't—so they're prioritizing a false idea of efficiency over fairness.

The scores Clifford and his colleagues were getting were meaningless—but the "bogus **WMD**" creating them was still gaining traction. The Obama administration sought to reform legislation that judged school districts based on test scores alone, creating a law that would let states turn around underperforming districts on their own terms. Even though strict correlation between a school's test scores and its overall health is falling out of favor politically, lawmakers and school board officials still aren't rejecting WMDs outright (or even recognizing that they're unfair). Value-added modeling, Clifford grimly predicts, isn't going anywhere anytime soon. Like inhumane scheduling software at major corporations, it's just too entrenched.

Once WMDs start influencing various sectors of society in significant ways, it's hard to eradicate them. Even federal legislation, at this point, is doing little to ensure that models and algorithms are performing the jobs they're supposed to. In many different fields, WMDs are too widespread to be toppled.

CHAPTER 8: COLLATERAL DAMAGE

In previous decades, local bankers controlled the money in any given town. People would suit up and pay a visit to their banker if they needed a new car, a mortgage, or a loan. The bankers were people and neighbors—they knew about a person's background and family in addition to having the numbers on their application form. In other words, the banker's judgement was human and thus biased. For millions of people, the human angle of banking was a liability: if they were poor, Black, or female, they might have trouble convincing a banker to give them a loan. When Earl Isaac and Bill Fair developed the FICO model to evaluate the risk of an individual defaulting on a loan, things seemed to be looking up—with an algorithm doing the work, there'd be no bias in the credit process.

Securing credit used to be a much more human process—and thus a much more biased one. Technology has evolved to make the process both fairer and more efficient. Because of FICO scoring, a person's credit history speaks for itself—and for many low-income people, non-white people, and women, that's an important step toward a more just world. But because of the tension between efficiency and fairness that O'Neil has described in previous chapters, it's clear that even a more objective process isn't necessarily a perfectly equitable one.

While FICO scores were relatively transparent, fair, and backed by consistently updated data, the use of scoring has changed significantly over the years. "E-scores" now aggregate everything from zip codes to internet behavior to purchase history to create arbitrary, unregulated, and unfair **WMDs**. Companies like Neustar and Capital One score credit-seekers lightning-fast using metrics like location and internet history to determine who's a worthy borrower. These e-scores create destructive rather than data-backed feedback loops. It's not clear what metrics they use to determine who will get a loan, and they're becoming more and more popular. By prioritizing efficiency over justice and transparency, they're becoming predatory and unfair.

FICO scores aren't WMDs—but e-scores, which are unregulated and widespread nowadays, certainly are. Their inner workings are mysterious, so while they deliver results faster, there's no telling whether the data they're using to decide people's futures is sound or not.

E-scores are taking society several steps backward from the fairness and transparency of the FICO scoring system. They're not looking at individuals; they're rating people in relation to a "blizzard of proxies." So, while e-scores don't do things like withhold credit from a Black lender based on race, they use things like zip codes—which can quite often be indicators of a person's race or class—to assess how similar people have behaved in the past, rather than how the person seeking credit has behaved in the past. There's no feedback that corrects these e-scoring systems when they make an error and arbitrarily place someone in the wrong category. And because the inner workings of e-scoring systems are hidden, no one can examine how they work in order to challenge or improve them.

E-scores are claiming to make the market fairer by using data that's more objective—but they perpetuate racism and classism all the same. The longer these scoring systems remain unregulated, the more they threaten to become the standard for lenders. And the more powerful and widespread they become, the more they'll deepen social divides by excluding minority and working-class people based on faulty proxy information rather than verified data.

Over time, creditworthiness has become a stand-in for virtues like a good work ethic and dependability, while bad credit signals "sins" that have nothing to do with being able to pay bills. Human resource management software now screens potential hires based on their credit reports, creating dangerous poverty cycles and feedback loops. "Framing debt as a moral issue," O'Neil suggests, is a huge mistake.

Hard data like credit scores are replacing the emotional intelligence and nuance that humans bring to hiring and other screening processes. In this way, technology is erasing humanity from the processes that define our lives, turning human errors and missteps into "moral issues" that stand in the way of efficiency.

The systems that crunch numbers and run data about people's lives aren't perfect—no-fly lists are rife with errors that keep ethnic and religious minorities from traveling safely and easily, while wealthy white people can pay for "trusted traveler" status and bypass security altogether. Credit report errors can make borrowing difficult or impossible for people with great credit. And scoring algorithms often mix up common names, meaning that having the same name as someone with a criminal history or poor credit can be a liability.

Even though people are being told that any error or misstep is a huge liability, the systems that judge them are far from perfect. This double standard means that computers can get away with more than people can, setting a dangerous precedent for a future that's increasingly governed by technology and data.

When an Arkansas woman named Catherine Taylor tried to secure federal housing assistance, she learned that her background check was full of mistakes and blended identities—she had many felonies on her report, some of which were tied to the alias of a woman named Chantel Taylor. Luckily, a housing authority employee helped Catherine find the errors and clean up the mess—but Catherine's case is evidence of a larger problem with how these systems are built.

Catherine Taylor suffered consequences—a smeared reputation and trouble getting the federal assistance she was owed—due to faulty data. It took human beings to help right Catherine's situation, whereas the technology that was judging her couldn't be trusted.

As algorithms become more automatic, more errors pile up in people's consumer profiles. These errors corrupt predictive models and give **WMDs** even more fuel. As computers become better able to learn from spoken language and images, these errors will only continue to pile up, creating uncountable instances of racist, unjust decisions made by faulty algorithms. These automatic systems now need a human hand to sift through their mistakes. Big Data needs to slow down and allow humans to play a greater role in sorting sensitive information, but the tech world is doubling down on predictive credit models.

This passage describes a dangerous domino effect that is already taking place within the tech world. Faulty data creates more faulty data—but broken algorithms aren't vetted or regulated intensely enough, so the systems become more and more error-ridden over time. Unless humanity takes on a greater role in checking these systems over and thoroughly inspecting them, they'll soon control more and more about how we live.

Facebook has recently patented a new type of credit rating based on social networks. For example, a young, white college graduate who spent five years volunteering in Africa might come home with no credit—but his connections on Facebook are successful and monied, and so he's able to get a loan. But a hardworking Black or Latino housecleaner from a poor neighborhood whose social networks might reflect "friends" who are unemployed or incarcerated will have a harder time securing financial help.

O'Neil illustrates how algorithms meant to make things fairer or level a socioeconomic playing field often end up encoding biases that already exist in society. They create racist, classist feedback loops that deepen social divides and keep the disadvantaged from equal access to certain opportunities.

Meanwhile, credit card companies like American Express have come under fire for revoking or lowering credit for customers who shopped at certain kinds of establishments. This plays into the idea that someone spending their money at Saks Fifth Avenue is more likely to pay off their card each month than someone frequenting Walmart.

By judging people based on their social networks or shopping habits, algorithms like the one described here make faulty associations. And when those associations lead to actions like denying someone a loan, a job, or a higher credit limit, they deepen social inequality.

Companies like ZestFinance, a start-up that calculates risk and offers payday loans at a discount, buy up data about their customers in order to inform how big of a loan they get, and what their interest rate will be. People are trading privacy for discounts—and if Big Data algorithms find something as minor as a spelling error in a mountain of data about an individual, it could affect their credit score.

WMDs have changed the economy in huge ways already—so much so that people are volunteering their private information in hopes of scoring a deal. Data is the foundation of the economy, so faulty or error-ridden data poses a major problem.

"Peer-to-peer" lenders like Lending Club, which hoped to become a "new kind of bank" when it launched in 2007, used a combination of credit reports and data to run their operations. These companies can analyze any data they choose to and develop their own e-scores and risk correlations without explaining the methodologies behind them. O'Neil suggests that compared to the systems in place today, the prejudiced loan officers and bankers of long ago don't look nearly as bad as they used to. At least borrowers, she writes, could "appeal to [their] humanity."

As e-scores grow in power, they're becoming less transparent and more capable of creating discriminatory feedback loops. So, while people might have faced judgement in the past, they could at least know what was counting against them. In today's Big Data economy, there's no transparency and no regulation, so it's har for people to understand how to navigate this confusing new realm.

CHAPTER 9: NO SAFE ZONE

In 1896, a German statistician named Frederick Hoffman who worked for the Prudential Life Insurance Company created a **WMD**. According to O'Neil, he published a 330-page report claiming that the lives of Black Americans were so precarious that "the entire race was uninsurable." Like many other WMDs, Hoffman's analysis was statistically flawed, racist, and unfortunately widespread.

WMDs, this passage illustrates, don't necessarily need to be tied to sophisticated technology or complex algorithms. Any time that math is used in a way that's difficult to understand and widely damaging, a WMD has been created.

For decades to come, insurers would cling to the idea that certain groups of people simply weren't worth insuring. Bankers and insurance companies would start delineating neighborhoods that they wouldn't invest in—this practice was called "redlining," and it wasn't outlawed until 1968. Yet redlining is still pervasive in U.S. society, and it's coded into contemporary WMDs that use flawed statistics to punish poor people and racial or ethnic minorities.

The racist redlining of Black Americans was, no doubt, a WMD that created widespread harm and deepened social divisions in the U.S. Though redlining has been banned for several decades, it continues to echo throughout U.S. society in other forms.

Many **WMDs** that perpetuate redlining are found in the insurance sector. Insurance grew out of the predictive field of actuarial science. In the late 1600s, mathematicians discovered that by comparing mortality rates of different people within a given community, they could calculate probable arcs of people's lives. Over the next several centuries, these predictions gave rise to the insurance business.

In the 1600s, math was used to create incredible predictions that had never been thought possible before. But in order to predict things like life expectancy, trends and similarities rather than individual circumstances became the metric by which people's worth were measured. Individuals were lumped together for efficiency's sake.

In today's world, more data about people's lives is available than ever before. Rather than making insurance predictions based on large groups, insurers are getting closer to being able to provide appropriate coverage based on the individual. Insurers use faulty proxies for responsible driving (like zip code and income) to create their own ratings, or e-scores—but because a lot of the information they use is based on credit and capital, insurance continues to work against the poor in many ways. Even drunk driving convictions count less in determining a person's premium than credit scores. By ripping off desperate, working-class people, these companies make a fortune off good drivers with bad credit scores. And because the factors that go into pricing at major insurers like Allstate aren't clear, their algorithms constitute **WMDs**.

Modern-day insurance continues to lump people together into certain categories using estimations and proxies rather than looking at an individual's unique circumstances. This perpetuates social inequality by discriminating against low-income people who may be good drivers or homeowners but have poor credit. The WMDs used to determine who's worthy of insurance aren't fair by any means—they're just efficient in terms of their ability to maximize insurers' profits.

In the age of Big Data, insurers can judge us by how we drive in entirely new ways. In 2015, the U.S.'s largest trucking company (Swift Transportation) started installing cameras in long-haul trucks—one pointed at the road, the other at the driver's face. The goal, according to Swift, was to reduce accidents—around 700 truckers die on the road in the U.S. each year. These fatal accidents are tragic, and they cost trucking companies a lot of insurance money (around $3.5 million per fatal crash). The additional surveillance also had another purpose, though: it let Swift gather a huge stream of data that could be used to optimize profits, compare individual drivers, and identify good performers.

Even though surveillance in truckers' cabs was ostensibly being used to make trucking safer, its true purpose was to avoid costly payouts for the trucking corporations. This illustrates how WMDs are often touted as tools that will make life safer and fairer for working people—when in reality, they're only used to maximize profits for companies. As such, this dynamic creates greater economic disparity.

Now, insurance companies offer regular drivers discounts if they agree to share their driving data through a small telemetric unit (like an airplane's black box) placed inside the vehicle. This has the potential to help drivers save money—especially younger drivers, who are often costly to insure—but it's also a big liability for poor or disadvantaged drivers. Driving through a bad neighborhood or providing evidence of a long commute each day might raise a driver's rate. Eventually, the insurance companies' promises to focus more on individuals becomes moot, because individual behavior is still being compared to that of others in similar demographics. While these systems are optional now, trackers, O'Neil asserts, will likely become the norm—and people will be punished for not having them rather than rewarded for consenting to them.

Again, O'Neil shows how, in the modern-day Big Data economy, more surveillance is often traded for a monetary break. So, working people who are desperate for insurance coverage, for instance, sacrifice their right to privacy in order to save money. This perpetuates economic disparity, and it sets a dangerous precedent for the future of surveillance and data-gathering techniques.

Insurance companies, O'Neil predicts, will soon start sorting people into new kinds of groups or "tribes" based on behavior. A decade ago, researchers at a data company called Sense Networks started to analyze cell phone data showing where people went. They could observe dots moving on maps to find similarities between groups of these dots. As the machine began sorting dots into different colors, only the machine knew what those colors meant—even Sense's cofounder admitted that human observers wouldn't be able to figure out what the "dots" had in common. This opacity, O'Neil asserts, is dangerous.

Major companies can now gather people's data with complete impunity. They don't even have to state what kind of data they're gathering, or for what purpose they're going to use that data. Data will continue to sort people, while most won't know how or why they're being sorted—in other words, humanity is now at the mercy of the assumptions that this advanced technology makes about them.

In 1943, because U.S. armies and industries needed every soldier or worker they could get, the Internal Revenue Service made a significant change: they gave tax-free status to employer-based health insurance. Within a decade, 65 percent of Americans were insured through their employers. This meant that employers gained a measure of control over their employees' bodies. Today, employers can offer rewards or impose penalties through "wellness" programs. These programs can create initiatives like "HealthPoints" in which employees accrue "points" by taking a certain number of steps in a day or going for a check-up. In other words, companies can penalize workers who don't consent to handing over data about their personal health.

Here, O'Neil shows how a measure taken in the name of efficiency—getting more people to enter the workforce at a crucial time—has slowly eroded privacy and allowed employers to get away with imposing judgment and bias on their employees.

Companies like Michelin have set employees goals for things like glucose, cholesterol, and waist size—employees who don't reach goals in at least three categories (out of several) must pay extra toward their health insurance. And in 2013, CVS announced that if employees didn't report their levels of body fat, blood sugar, blood pressure, and cholesterol, they'd have to pay $600 a year. This drew public ire, since the company used BMI (or body mass index) as a measurement of health—but BMI scores are "crude numerical prox[ies]" that were originally based around "average" male body types. In other words, their usefulness has been all but debunked.

The same way that nefarious e-scoring programs use proxy data like zip codes to determine who's worthy of being insured, companies are now able to employ proxies like BMI to withhold or grant privileges to their employees. This widens inequality because BMI scores don't necessarily give an accurate picture of a person's health, so people with high BMIs may be unfairly penalized. In this way, employers essentially have permission to treat their employees unfairly based on outdated, flawed data.

Even though companies assert that they're taking these invasive measures in the name of health, wellness programs don't lead to lower healthcare spending—and there's no evidence that they make workers healthier. O'Neil asserts that wellness programs aren't yet full **WMDs**, since they're often quite transparent. But they do show that employers are "overdosing" on employee data, trying to score potential workers and predict their productivity. If companies start creating their own health and productivity models, O'Neil suggests, the industry could very well become a full-fledged WMD.

Because these initiatives that pry into employees' personal health information are geared toward providing employers with personal data about their workers, O'Neil implies that they're classist and harmful. Without regulation, they could become the norm, and people's personal freedoms will continue to erode.

CHAPTER 10: THE TARGETED CITIZEN

O'Neil imagines creating a petition for tougher regulations on **WMDs** and posting it to Facebook. As soon as she hits "send" on the post, the petition belongs to Facebook. The site's algorithm gets to decide what to do with it and whom to show it to, based on the data it has about each of O'Neil's "friends." For many of them who don't engage with many posts and who never circulate petitions, it'll get buried in their feeds—but for others, it'll pop up at the top. Facebook isn't the "modern town square" it might seem to be. Its powerful algorithms and news-molding infrastructure have allowed it, in many ways, to "game" the U.S. political system.

By illustrating how Facebook's selective algorithms work, O'Neil shows how models are eroding transparency. Facebook isn't a "town square" where people can gather to, openly discuss the same information. Rather, because Facebook hides certain things from certain people, it's creating an uneven and contentious media landscape. And this is dangerous when Facebook applies its algorithm to news, because it can influence people politically.

During the 2010 and 2012 U.S. elections, Facebook created experiments to hone a tool called the "voter megaphone" that would allow people to spread the word about voting. Facebook was encouraging over 61 million American users to get out and vote by leveraging peer pressure against them. At the same time, they were studying how different types of updates influenced voting behavior. Because the profits of companies like Facebook, Google, Apple, Microsoft, Amazon, and Verizon are heavily regulated by government policies, these companies often spend a lot of money lobbying and donating to the political system. Now, they can influence Americans' political behavior and, as a result, the shape of American government.

In today's world, technology and data companies have a say in U.S government. Now that these companies hold political sway, they're coming up with algorithms and models that will more clearly show them just how much influence they possess—and how directly they can change how people participate in civic life.

Facebook's grand experiment with the voter megaphone showed that they increased turnout by nearly 350,000 people—a big enough group to swing whole states. Facebook used the initial 2010 experiment to study how our friends' behavior impacts our own—and in 2012, they took things a step further. For two million politically engaged users, researchers tweaked the algorithm to show these people more news instead of social updates or funny videos. Researchers wanted to see if getting news from friends would change people's political behavior—and it did. Voter participation in the group went up by three percent.

Algorithms can change the way people see the world—and thus how they participate in society. People will respond to information they receive from outlets they trust, and Facebook's experiment showed how powerful algorithms are in manipulating this trust. Although increasing voter turnout is a positive result in theory, O'Neil implies that readers should be alarmed about the fact that tech companies have the ability to change outcomes in elections.

Editors at news outlets, of course, decide what their readers see, and from what angle they see it. But when a news outlet covers a story, everyone can see it. When Facebook delivers news, the process is mysterious—it's not the "neutral go-between" it might seem to be. Most users are unaware that the company is tinkering with their news feeds and filtering what they see (and which of their posts their network can see). Research indicates that showing Facebook users positive updates puts them in better moods, while showing them negative updates from friends puts them in worse ones. Emotional states can be transferred through the internet—and through Facebook's algorithms.

Again, O'Neil is using Facebook's information delivery model as an example of how powerful algorithms can be. These models have the power to change people's opinions and even their moods. This poses a problem because Facebook's algorithms are hidden from the public, so users are being manipulated without their consent.

O'Neil doesn't believe that Facebook's researchers are actively trying to game the political system. But she does believe that Facebook has the power to determine what people learn about the political system, how they feel about it, and how they participate in civic life as a result. Facebook and Google haven't yet turned their algorithms into political **WMDs**, but the "potential for abuse is vast."

Facebook might not be actively trying to change how its users think and feel, especially where politics are concerned. Nevertheless, Facebook's algorithms do have enormous influence over their users—so the company needs to make sure it doesn't misuse that influence.

In the spring of 2012, Mitt Romney had all but secured the Republican presidential nomination. He traveled to Florida for a fundraiser at the home of Marc Leder, an investor who'd given over $300,000 to Romney's campaign. Romney assumed he was walking into a closed setting with a likeminded group of people—but as he let loose during his speech with traditional Republican talking points, he underestimated his audience. One of the caterers captured Romney claiming that 47 percent of the U.S. population were nothing but "takers," posted it to the internet, and let the world see.

This passage is an example of a real-life model in action. Romney had developed a model for how he'd speak to attendees at Leder's gathering—but he'd failed to account for a crucial piece of data (the demographics of everyone in attendance), and so his model backfired.

Most politicians tailor their pitches for lots of different subgroups on the campaign trail. This is, essentially, a form of modern consumer marketing that's driven by carefully gathered data. In the incident at Leder's house, Romney was speaking based on one set of data—but he ignored that other groups might be in attendance.

Politicians can use data about potential voters to help them get ahead, but doing so can also hinder their campaigns. This illustrates how crucial it is for data to be thoroughly vetted before it's used to model an outcome—especially when that data is related to politics.

Nowadays, Big Data has given politicians lots of powerful tools for targeting "micro-groups" of citizens for votes and donations through carefully honed messages. In July of 2011, the Obama campaign started hiring analytics experts who would help create and target groups of like-minded voters. Rayid Ghani, one of the campaign's data scientists, had previously worked on projects for a consulting that analyzed grocery stores' consumer data. This information was used to create customized shopping plans for many kinds of shoppers: coupon-clippers, brand loyalists, foodies, and so on. Now, Ghani was trying to see if similar calculations would work on swing voters (those who aren't firm supporters of any one candidate or political party).

The Big Data economy isn't just changing how companies work with their employees, how colleges handle admissions, or how insurers decide whom they should cover. It's changing political and democratic institutions all over the world, as well. Algorithms play a significant role in contemporary politics, because faulty data can now quite literally decide the fate of an entire country. This is a dangerous precedent to set as tech becomes a bigger part of everyday life with each passing year.

After conducting deep interviews with a few thousand people from different groups, Rayid and his team set out to find voters who resembled them by sifting through the data and demographics of the people they'd interviewed and creating mathematical profiles of them. Then they scoured national databases to find people with the same profiles so that they could target each group with advertisements to see what metrics appealed to them. By the end of all these tests, the researchers had their group of 15 million swing voters.

By finding similar people based on users' voluntarily submitted data, Rayid's team was essentially creating proxies. From this, they were able to reach a massive audience just by extrapolating information about a small group, whether that information was relevant to everyone in their target audience or not. This illustrates how readily available—and yet potentially useless—proxy data is.

Four years later, Hillary Clinton's campaign would build on the Obama research team's methodology to create a data system that would let them manage millions of voters. But many of the methods used to gather data that campaigns use aren't necessarily legitimate. Data firms keep tabs on users' "likes" and use those to rank them on the scale of "big five" personality traits (openness, conscientiousness, extroversion, agreeableness, and neuroticism). Then they develop targeted ads based on this information. Not all of these methods are useful—but some are, and they've essentially turned the voting public into a kind of fluctuating financial market.

Political campaigns are now treating voters like a tech company might treat its users—by appealing to their lowest common denominator. This, in O'Neil's estimation, is a dangerous new development. Political campaigns should be held to a higher standard than leisure pursuits like social media platforms, because they're deeply intertwined with the U.S.'s democratic process.

Lobbyists and interest groups, too, use these microtargeting tactics to reach more people. This can be dangerous—certain political groups can create targeted advertisements full of false information that spread rapidly. The anti-Obama "birther" campaign got off the ground this way, as has a lot of misinformation about things like abortion and immigration. Even television is moving toward personalized or microtargeted advertising—and as individualized ads become more common, it'll be harder to understand or even access what our neighbors and friends are seeing. Political marketers have scores of information about us, but we have little about them or their methods.

Microtargeting allows political campaigns to reach more voters—but it also means that the approach behind microtargeting can be used to spread false information. Marketers from all sectors of the economy know that appealing to people based on data-backed research is a good way to bring in new users, new customers, or new voters. But the algorithms that enable that efficient approach are also being used to undermine facts and democracy.

Microtargeting is vast, largely hidden, and unaccountable or unregulated—so it is, in O'Neil's estimation, a **WMD**. And it's actively undermining and threatening U.S. democracy. Additionally, what's so frightening about political microtargeting is that it's not aimed only at the rich or the poor—it's aimed at everyone except for those who aren't expected to vote. This creates a feedback loop that keeps uncertain or disenfranchised voters out of the civic system.

Like many other WMDs, political microtargeting could be used to help and educate people—but instead, it's used to manipulate and mislead them. O'Neil implies that humanity needs to swiftly and carefully begin regulating WMDs so that they won't be used to threaten the truth and destabilize democracy in the U.S. (or abroad).

CONCLUSION

WMDs cause destruction and chaos throughout society: in public schools, colleges, courts, workplaces, voting booths, and more. But it's too late to disarm these weapons one by one—they all feed into one another. Data encourages companies to send people predatory ads. It also encourages police to go into vulnerable neighborhoods, and then it influences the courts to give the people whom police arrest longer prison sentences. All this data tells other WMDs that these people are high risks—so they're blocked from jobs and watch helplessly as their interest and insurance rates ratchet up. All the while, WMDs keep the wealthy and comfortable in silos of their own, ignorant of others' suffering. WMDs are part of a "silent war."

Here, O'Neil walks readers through the interconnected nature of the processes behind applying to schools and jobs and seeking credit and insurance. In doing so, she's illustrating the reason that WMDs are so dangerous: because the data used to create them now feeds multiple systems at once. It's hard to disengage from WMDs because they're virtually everywhere.

Corporations can right wrongs in their algorithms—for instance, even though President Bill Clinton signed the Defense of Marriage Act into law in 1996, IBM promised a week later to extend benefits to its employees' same-sex partners. They did so not necessarily because of morality, but because other tech giants were already doing so, and they didn't want to lose employees to competitors. So, in a bid to attract a growing talent pool of LGBT workers, IBM corrected an unfairness.

Here, O'Neil shows how a simple change in policy at a crucial moment had a ripple effect. Even though IBM made this particular move to maximize their efficiency and competitiveness in the market (but did so in the name of fairness), there's still room for major corporations to make big changes in the name of equity and objectivity.

In that scenario, everyone won—but companies aren't always so incentivized to dismantle their **WMDs**. Many WMD victims are the most voiceless and disenfranchised: the poor, the incarcerated, the vulnerable. These easy targets are where all WMDs start operating. But it won't be long, O'Neil predicts, before they evolve and spread, targeting the middle and upper classes as they search for new opportunities.

Just because predatory college loans or sky-high insurance rates are targeted at working-class people, it doesn't mean that corporations won't start aiming their WMDs at other groups. If we don't stand up for the vulnerable, O'Neil is saying, soon we'll all be victims of predatory WMDs.

The main difference between the **WMDs** of the present and the prejudiced human errors of the past is simple: humans can evolve, learn, and adapt. But automated systems are stuck in time—engineers have to change them as society progresses. So essentially, "Big Data processes codify the past" rather than inventing the future. Only humans have the "moral imagination" needed to create a better world. O'Neil asserts that humanity is in the throes of a new kind of industrial revolution—and it is urgent that we learn from the mistakes of the last one, which exploited workers and endangered lives in the name of profit.

Even though algorithms and machines can "learn" in a sense, they're simply not human—they can't imagine things, they don't have a moral compass, and they are only as good as their creators. But humanity has the capacity to understand the stakes of its present moment—and so humans, not models, should be the ones in charge of humanity's most important processes and questions.

We need to regulate the mathematical models that increasingly run our lives, and we must start with the modelers themselves. Like doctors who swear to the Hippocratic Oath before obtaining their medical licenses, O'Neil suggests, data scientists need to abide by certain moral codes and strictures that prevent them from doing harm to others. Regulating **WMDs** would be difficult and deeply involved—but O'Neil argues that even if it comes at a cost to efficiency, we must start to "impose human values" on WMDs and "get a grip on our techno-utopia."

Right now, the tech world and the Big Data economy are uncharted territories in many ways. Those in charge are looking to maximize profit and influence as wide of an audience as possible—and there need to be some checks and balances in place. Without human regulation on these incredibly powerful technological tools, society will become more stratified, democracy will come under threat, and our "utopia" may soon devolve into a dystopia at the hands of predatory technology.

In order to disarm **WMDs**, we must admit that they can't do everything. We must measure their impact by auditing their hidden algorithms and studying their biases and shortcomings. Unfair systems, like the value-added model used to score teachers, must be dispatched entirely. Rather than letting negative feedback loops slip through the cracks, analysts must figure out how WMDs can create positive feedback loops that change lives and benefit society. While some algorithms, like Amazon and Netflix's, should be allowed to sort the kinds of entertainment people enjoy, recidivism models and other algorithms used in the justice system must be held to unimpeachable standards—even if it means revising them and changing their inputs altogether.

Here, O'Neil suggests that there should be different standards for algorithms in different sectors of modern life. There are models that are relatively simplistic and that don't require much oversight, like the algorithms on streaming platforms that suggest programming based on what users have watched before. But when it comes to the education, the justice system, and politics, there needs to oversight and regulation. This is because if data is used the wrong way in these arenas, it could threaten social stability on a large scale.

Not all potential **WMDs** are nefarious. But the point is that we need analysists and auditors to maintain the systems that govern our lives and make them more transparent. Internal audits alone aren't enough, O'Neil states, because companies that examine their own algorithms can't be held accountable. Outside input is needed to make sure that companies like Google and Facebook stay in line. And regulations and transparency are needed in peer-to-peer lending, healthcare and health insurance, and credit score models.

The technology to make the Big Data economy more transparent exists—we just need to start using it. Even though data is being used in legitimate and transparent ways in many sectors, the potential for harm is too great to allow technology to influence our lives without any protective measures in place.

In 2013, O'Neil began working as an intern at New York City's Housing and Human Services Departments—she wanted to build the opposite of a **WMD**, a model that would help stop houseless people from getting pushed back into shelters and help them finding stable housing. Her and her team's research found that families who received Section 8 affordable housing vouchers were less likely to return to shelters. But the city government was trying to get families away from Section 8 to a new program, Advantage, that limited subsidies to a few years. Public officials were not happy to hear about O'Neil's team's research.

O'Neil's experience with a branch of the New York City government shows that even public service organizations are using technology and data to exploit people rather than helping them prosper. The NYC Housing and Human Services Departments were using technology to prey on people's desperation and deepen class divides rather than using the data they'd gathered to actually change lives.

But Big Data, O'Neil asserts, should be disruptive when it comes to things that actually matter, like human rights. There are so many mathematical models out there today, O'Neil writes, that could be used to do good—but instead, they often wind up being abused. Yet there's hope in the form of supply chain models that seek out sweatshops and other places where slave labor is being used to build products, and predictive models that try to pinpoint houses where children are more likely to suffer abuse.

The models O'Neil describes here have the potential to really help people—but if placed in the wrong hands and stripped of their transparency, they could easily become WMDs. She's underscoring the importance of using innovative technology to help people rather than exploit them.

O'Neil hopes the **WMDs** that are around today will soon become relics of the past. She hopes that we can learn from our present moment—the early days of a "new revolution"—and learn to bring transparency, fairness, and accountability to the age of Big Data.

Humanity is indeed in the midst of a "new revolution"—and the Big Data economy offers lots of opportunities for social change, economic reform, and a more egalitarian society. But if used incorrectly, WMDs could actually erode democracy and create more social division. So, humanity needs to recognize the weight of our present moment and rigorously ensure that algorithms and the technology they power are objective, fair, and reliable.

AFTERWORD

In 2016, millions of people were shocked when news outlets' algorithms failed to accurately predict the results of the contentious U.S. presidential election. Because feedback loops for midterm and presidential election models aren't updated so frequently, there's a lot of room for things to change quickly: and indeed, a lot of things changed between 2012 and 2016. A rise of populist politics, media skepticism, and people's reluctance to contribute data to polls meant that the predictive models' results were impossibly skewed by biases and misinformation. The narrow gap between the candidates, Hillary Clinton and Donald Trump, wasn't as narrow as it appeared to many people.

Faulty polling data, O'Neil shows here, likely had a big impact on a major U.S. election. People trusted the polls' predictive power—and they voted (or didn't vote) because of what the polls signaled. Polling is directly tied to civic life and democracy in the U.S.—so the algorithms and models that govern it need to be thoroughly vetted and held to high standards in order to ensure accuracy and transparency.

While political polls are influential and somewhat mysterious, they're not necessarily destructive—so they're not quite **WMDs**. But because people gave polls so much power in the 2016 election only to see them completely miss, O'Neil is hopeful that they'll be given less and less power in politics as time goes on.

Polls can create the illusion of being able to predict the future. But because they're not always accurate, their weight in our modern political system is potentially dangerous. Polling should become more transparent—and the data they gather should be vetted more thoroughly.

Polling wasn't the only algorithmic failure in the 2016 election season. Facebook's "Trending Topics" algorithm, which was meant to eliminate news bias, ended up performing erratically and flooding the site with "fake news" and other kinds of misinformation. O'Neil suggests that wonky algorithms like this one shouldn't necessarily be banned or dismantled forever—but there must be algorithmic accountability that starts and ends with the developers themselves. Journalists are increasingly working to help people understand how algorithms perpetuate bias or create discrimination, hoping to make them easier to see through and understand.

When WMDs begin to spiral out of control, it's up to humanity to call out their destructive nature. If tech companies won't make their algorithms transparent, then everyday people and arbiters of truth, like journalists, need to encourage transparency in other ways.

O'Neil isn't sure whether there will ever be a simple, widely recognized definition of what makes an algorithm fair, but she's grateful that people are finally discussing what that definition might look like. By continuing to set standards for algorithmic accountability, she suggests, both technology and contemporary ethics will improve.

Here, O'Neil's essentially saying that the desire for technological progress doesn't outweigh the need for fairness and transparency. Even though WMDs operate in many different sectors of contemporary life, there's still hope that by putting a human hand back into tech, reforms can happen.

By weighing harms instead of squabbling over fairness, O'Neil suggests, we can dismantle **WMDs** slowly but surely. With every algorithm created—for example, one that seeks to determine which households are most likely to be hotspots of child abuse—there are harms associated with both false positives and false negatives. Determining which is the greater harm is difficult but necessary work, and it will help ensure that algorithms are functioning the way they're supposed to.

By making sure that we see and understand the data that's used to create WMDs, we can stop faulty or irresponsible algorithms from ruining lives. Technology can be useful and transformative—but without examining its potential for confusion or even harm from all angles, humans are letting faulty technology take over too many aspects of modern life.

Artificial intelligence algorithms cannot distinguish between the truth and lies—so asking Google "who won the popular vote" in the 2016 election won't always yield accurate results. Instead, the results might be contaminated by conspiracy theories. Data scientists, then, need to work to ensure that the data these algorithms use represents the world as it is.

Because the very concept of objectivity and truth is at stake, we must interrogate how data is collected and processed. This is a difficult task, to be sure, but one that O'Neil suggests is worth the effort to prevent the proliferation of WMDs.

"Truth" can look like different things from different points of view—and even mathematical proofs can be full of mistakes. But Big Data has a duty to clarify the noise—not contribute to it even more. Big tech companies exert a huge amount of control over contemporary society because they control people's data. And if that data remains privately owned and used, the algorithms it creates can't be trusted.

When tech companies, the algorithms they use, and the kinds of data they collect aren't regulated or vetted, society actually becomes less *efficient. Humanity needs to step in and thoroughly examine the role technology plays in our lives. Otherwise, we will likely be victimized by error-ridden algorithms that deepen social divides without even delivering on their promises of efficiency and fairness.*

Algorithms aren't going anywhere—if anything, they're only going to become more common as time goes on. In light of that fact, O'Neil argues, it's time to hold algorithms accountable in the long term by making sure that they are legal, fair, factual, and capable of change. We must focus our efforts on improving how algorithms work, O'Neil warns, because "we can't afford to do otherwise."

WMDs are a threat to personal privacy, protections for minority groups, and indeed democracy as we know it. By failing to regulate how companies collect and use data about us, humanity is setting a dangerous precedent of living at the mercy of new technologies that seek to maximize profit and efficiency at any cost.

HOW TO CITE

To cite this LitChart:

MLA

Tanner, Alexandra. "*Weapons of Math Destruction.*" *LitCharts*. LitCharts LLC, 16 Jun 2021. Web. 16 Jun 2021.

CHICAGO MANUAL

Tanner, Alexandra. "*Weapons of Math Destruction.*" LitCharts LLC, June 16, 2021. Retrieved June 16, 2021. https://www.litcharts.com/lit/weapons-of-math-destruction.

To cite any of the quotes from *Weapons of Math Destruction* covered in the Quotes section of this LitChart:

MLA

O'Neil, Cathy. Weapons of Math Destruction. Broadway. 2017.

CHICAGO MANUAL

O'Neil, Cathy. Weapons of Math Destruction. New York: Broadway. 2017.