Universidad Pontificia Comillas ICADE

BIG DATA AND BUSINESS ETHICS

Gabriela Muñoz Fernandez

MADRID, JUNIO 2022

ABSTRACT

We live in an era where algorithms are the norm rather than the exception. Automated systems are increasingly making judgments about our lives, such as where we go to university or whether we can get a promotion or a loan. Theoretically, having everyone judged by the same standards should lead to a more equitable society. Research shows that today's models are uncontrolled and unquestioned no matter how inaccurate they are. More worrisome, they perpetuate discrimination, boosting those who are fortunate at the expense of those who are less fortunate and eroding our democratic system. It demonstrates the shadowy side of Big Data, which has been condemned for violating people's personal information. Large, complicated data sets and new assumptions have been used to target individuals for things they didn't realize they wanted, to alert family members that someone is pregnant, to determine what news people see and to charge consumers more depending on their buying habits.

Key words: Big Data, algorithm, mathematical models, IoT, artificial intelligence, Silicon Valley.

INDEX

1. INTRODUCTION

1.1. Objectives

This investigation paper has the objective of serving as an educational guide to explain how algorithms, artificial intelligence, and machine learning work in the 21st century, how they infiltrate into our everyday lives, and to combat secrecy with clarity. Thereby, it will try to answer the following question:

> *Have we eliminated human bias thanks to Big Data or camouflaged it in algorithm models?*

1.2. Methodology

This study will be first, descriptive, of Big Data and its components. Second, search how algorithms are biased and maybe not as neutral as one perceives them, and most importantly, find out how they govern our everyday lives and how this intrusion is only going to increase in the future. And third, to provide real life examples of people who have been affected by this industry. Big Data is presented as an "Industry not as a technology" (Martin, 2015) and it is imperative for leaders and developers to set ethical guidelines as it keeps growing. All of the data will be collected from different research studies and articles in this field.

1.3. Motivation

We always see the advancements of technology displayed on the news and the benefits they suppose for humans and societal development. However, there's another story Silicon Valley advocates won't tell: the dark side. Big Data is very complex and hard to understand, but because it has clawed into nearly every aspect of our daily lives, it is imperative to understand it and help others gain accessible and understandable knowledge of how it works, expose its flaws on how information is actually used so that we can open our eyes on how disturbing and deeply important it is to change the course of how the Big Data Industry is moving towards, before it's too late.

The following research paper tries to make the unknown known, to shed light on the fact that our reliance on algorithms has gone too far, and to serve as a subtle reminder that Big Data is only as good as the people who use it, and despite the technical complexity of the subject, to assist those who lack the mathematical skills to understand the whole picture for better critical thinking in the age of Big Data.

2. WHAT IS BIG DATA?

The big data economy was developed by mathematicians and statisticians who studied human preferences, movements, and purchasing power. Their models might forecast our trustworthiness and assess our potential as thieves, students, or employees. Big data promised dramatic results since it was a computer capable of rapidly processing millions of documents, such as loan applications, and sorting them into lists ranked from best to worst. The fairness and equality of these models were classified as such because we all believed that they were free of human preconceptions. Nevertheless, these models had already incorporated human biases into the software systems. All but the most knowledgeable among us had no idea how these models functioned. No one could argue with the decisions, which tended to favor the wealthy while penalizing those who were already well off.

Why do they punish the oppressed? Models are designed to analyze vast numbers of people because doing it one by one would be extremely time-consuming. For example, entry-level positions are commonly filtered by a computer system while higher-ranking positions are most likely filtered by a human; while the wealthy are handled by humans, the masses are handled by machines. This is why it is so important to understand the algorithm, as it governs the results of these mathematical models, and upon these results, all the decisions are made. However, the models are of an “opaque” nature, where only a few can understand the algorithm, keeping the public in the dark, as they can’t question the model’s results. If one of these models determines you are not a good professional and you get fired because of this, what can you do? Even if you are able to understand their inner workings, it's impossible to make a case against a math model; that's part of what makes them so terrifying, and if you are a human victim of these judgements, you are subject to stricter evidentiary standards against these algorithms.

3. WHAT IS A MODEL?

A model is a simplified depiction of reality that can be readily comprehended and from which key facts and actions may be deduced and applied. Thanks to historical data available, we can analyze current situations and possible outcomes. The goal of a model is to run a variety of alternative scenarios at each point in order to find the best possible outcomes. In the book Weapons of Math Destruction by Cathy O’Neil they use the example of Michael Lewis 2003 bestseller, Moneyball, to depict a good model. They explain how they measure if to change a player for another, who would give the biggest chance of success; this is what a model can provide, all the outcomes varying the variables. In the beginning, statisticians began to study what all of the player’s historical data meant (each player's home-runs, how many strikes and hits, etc.), how they might translate it into victories, and how management could optimize performance with the minimal investment possible. All of this data is accessible to the public and easy to comprehend (we all understand what it means when a player hits the ball). They accumulate an enormous amount of historical data related to the results they are trying to forecast, and there is a constant flow of new statistics coming in (new players). The models for baseball reflect fair ones; they are in a constant exchange of information with the results of each season and the forecast they are attempting. Thus, the models are constantly updated. However, most of the time, the data used for models isn't as dynamic, the models aren't in a constant communication with the real-world's data and neither sustain relevant with historical data, but rather use proxies, as most of the time we don't have the exact data necessary for our model because it doesn't exist, and thereby enter into a "toxic feedback loop" (O'Neil, 2016), which we'll see reflected in several real-life cases that use predictive models, like in risk assessment tools, college rankings, personalized advertisements, Facebook, personality tests, and credit scores. Not only might they not have the best data to begin with, but they are also programmed by humans, and despite their perceived objectivity, models actually reflect the aims and ideologies of their creators. O'Neil emphasizes in her book how mathematical models have their own set of beliefs incorporated into them and calls them "weapons of math destruction (WMDs)". Despite the example of the *just* models used by professional baseball leagues, we see numerous examples where models determine who gets fired, who gets a loan, or who serves more time in jail, and most of the time the outcomes harm those who need help most. Furthermore, when evaluating a model's efficiency, we must also examine whether the model works or not, not simply for who designated it but for what objective or what the organization is attempting to achieve. The question is whether we have eradicated human bias or merely disguised it with technology. The models are complex and mathematical in nature. However, host assumptions, some of which are biased, are embedded throughout these models. And the difficulty is that the inner workings of a model are hidden in algorithms, accessible only to a small group of people. However, many businesses go to great lengths to conceal their models' output or even their very existence. Some argue that the algorithm serves as their "special formula," which is essential to their business model. Licensed lawyers and lobbyists may be used if necessary to protect intellectual property. These custom algorithms themselves are worth millions of dollars to the likes of Amazon, Facebook, and Google. But as we'll see, these models are intentionally opaque black containers. That complicates the question of whether or not it is fair.

Damage, opacity, and scale are the three components of an unfair model. We'll see all of them in some form or another in the cases we discuss. If, for example, the recidivism scores are not completely opaque, you could claim that in certain situations, inmates are able to view the scores that are generated. However, they are full of mystery, as the prisoners are unable to see how their response led to their score. The scoring formula is kept secret. It may appear that just a few models meet the requirement of scale, as at this point, they aren't quite as big as they could be. However, they represent a potentially hazardous organism that is poised for rapid expansion. In addition, algorithms are capable of hopping from one field to the next, which is exactly what they frequently do. Finally, it's worth noting that not all of these things are harmful in the same way. Which is not to say that all of their efforts don't benefit certain individuals. Regardless of whether some individuals profit, that is not the purpose. It's because there are so many people who are hurting. Algorithms in these models close doors in the faces of thousands of individuals, often for the most tenuous of justifications, and there is no appeal. They're treating you unfairly.

4. BIG DATA INDUSTRY'S SUPPLY CHAIN

As Kirsten E. Martin explains in Ethical issues in the Big Data Industry, there are issues in the supply chain of this industry upstream and downstream. Looking at Big Data's supply change, what is being passed from one firm into the other is information, whereas in the most traditional industries, it is raw material that is being passed and transformed. On the one hand, the downstream applications of Big Data might be viewed as creating both helpful and dubious results. However, the risks associated with Big Data use should outweigh the benefits. Nonetheless, selling data increases the possibility of secondary misuse, which could have a negative impact on consumers. While there is a risk of injury from wrong information downstream in the supply chain, there is also a risk of harm from corrected inferences. She summarizes the negative consequences of downstream as: "value destruction for shareholders, diminished rights for stakeholders and disrespectful to someone involved in the process" (Martin, 2015) which, in retrospect, are also effects of data provided by upstream information, she concludes. For example, you can be denied credit, lose your job, have your secrets revealed to your family, or have to pay extra for your insurance as a result of downstream information. Another consequence, explained by Ryan Calo, is that firms can affect customers on a personal level and provoke vulnerability in their marketing when they have access to more information about consumers and the ability to fine-tune the consumer experience (Ryan Carlo, 2014). Martin, raising a good question, further than if there is value creation or destruction for users, whether or not individuals' rights are respected during the data-use process. On the basis of pre-existing data, Big Data may generate learnt prejudice algorithms, Martin concludes. Learned prejudice reinforces previously established prejudice by building predictive algorithms on historical data patterns. Such algorithms may result in undesirable effects, such as discrimination, whether inadvertent or purposeful (Martin, 2015). Even if value is not annihilated, people might be insulted by being reduced to a simple category, such as: "alcoholics, erectile dysfunction sufferers and even as *daughter killed in car crash*" (Martin, 2015), specifically, if such a category is significant in its own right, such as being the *victim of a crime*. On the other hand, there are other consequences in the upstream supply chain of Big Data: a major problem for companies in the data supply chain is the risk of working with untrustworthy data providers. The ability to get more and more data from more and more sources means that you need big, complicated, and spread out data sets from a lot of different places, such places: "include consumers, products, location, machines and transactions" (Martin, 2015). Upstream data sources may be problematic in the Big Data industry because of the: "quality of information, biases in the data and privacy issues " (Martin, 2015). Data quality refers to the method of acquiring such data; data bias refers to data being skewed towards a user category; and data privacy as suppliers should make sure they are not disrespecting the privacy of the users they are collecting the data from. An ethical problem Martin raises within the Big Data Industry is that many practices are becoming generalized and used by all, which results in systematic issues.

5. RISK ASSESSMENT TOOLS

In most U.S. states, judges use risk assessment tools to determine the sentences of criminal defendants. These assessments are algorithms that use variables and probabilities on different factors, such as age, criminal record, family history, employment records, etc., to predict the likelihood of recidivism. Critics have raised a number of questions, such as: “Is it fair to make decisions in an individual case based on what similar offenders have done in the past? Is it acceptable to use characteristics that might be associated with race or socioeconomic status, such as the criminal record of a person’s parents? And even if states can resolve such philosophical questions, there are also practical ones: What to do about unreliable data?” (Barry-Jester, 2015). Nowadays, the lengths of sentences do not only depend on past and present actions, like what crimes did you commit in the past or what crimes did you just commit, but also, future actions, like what is the probability of committing crimes again. The algorithms in risk assessment tools depend more and more on Big Data. Risk assessment tools used to “simply add up points from a checklist” (Humerick, 2020) whereas now that they rely on Big Data, these algorithms are automated and recollect historical data that is publicly available to anticipate future criminal activity. Professor Aziz Hug from the Law School at the University of Chicago has defined this trend as: “ algorithmic criminal justice is "the application of an automated protocol to a large volume of data"” (Humerick, 2020) where they may draw inferences about criminal behavior based on a sample size. Therefore, the question that this raises is why a computer is the one deciding whether or not to send a person to prison. The latter question is illustrated in the case State v. Loomis, where a risk assessment tool called COMPAS was used. Eric Loomis was identified by COMPAS as a high-risk individual, and he appealed that COMPASS violated his rights. An expert declared at court that: “[t]he Court does not know how the

COMPAS compares that individual's history with the population that it's comparing them with. The Court doesn't even know whether that population is a Wisconsin population, a New York population, a California population . . . ” (Humerick, 2020) and this is partially due to the fact that these tools assessments are privately owned and don’t disclose the algorithm. While the case was taking place, in 2016, ProPublica, a non-profit journalism company, alleged that a risk assessment called COMPAS was biased towards black people: “the algorithm incorrectly predicted black defendants to be high risk more often than white defendants, while also incorrectly predicting white defendants to be lower risk more often than black defendants” (Humerick, 2020), while COMPAS succeed on predictions of recidivism, it failed on the rates for this in respect to race. To conclude, because of the condition of crime and police in the United States, a criminal sentencing algorithm cannot promise both predictive accuracy and equalized chances, which is why the COMPAS algorithm has been deemed unconstitutional. It is necessary to be "unfair" to certain defendants in order to be "fair" to others for COMPAS and its kind. Because race is a protected class under the Constitution, neither the federal government nor the states can make decisions based on race, thereby, no risk assessment tool expressly adds race, despite the fact that many of these algorithms account for the same reoffending variables and personal data. Adding race to the algorithm could fix its bias, and some organizations have started using proxies to race. Yet, we see a trend increasing in the use of these tools, which could be due to the fact that: “Judges, legislators, and even private companies themselves can deflect blame onto an algorithm for any mistake, rather than justify their own decisions. As such, it is highly unlikely that we will see the elimination of risk assessments at the state or federal level anytime soon” (Humerick, 2020) and the solutions to this are also hard to execute.

An algorithm created by humans, based on prior data that has been implicitly imbued with decades of racial bias from American culture, can reflect and strengthen contemporary society's preconceptions. Humerick proposes developing a “race-conscious risk assessment” algorithm for criminal sentencing as the best course of action, which is supported by the work of experts in the area.

6. COLLEGE RANKING

In 1983, *U.S. News and World Report*, a second-tier magazine that was struggling to get readers, decided to issue a college ranking, which had never been done before. Before college “rankings" people decided where to go to university based upon anecdotes of who went to what university and was successful and word of mouth. What university was better was subjective, it was a personal experience. However, when the new issue launched, the ranking was based on: “the results of opinion surveys it sent to university presidents ” (O’Neil, 2016) and every university complained back to the magazine, furious and questioning their ranking metrics. Therefore, *U.S. News and World Report* decided to base their ranking upon a model. Now, the reporters had to decide what variables would go into deciding what matters most in an education and how much weight would be placed on each; they would also consider a small portion into the model of college officials' own opinions from all throughout the country. In 1988, the first data driven rating was published, and now every university has their reputation on the line because of this model. If your university ranked low in the 15 areas that encompassed the model, a vicious cycle would start where top students and professors would avoid you, alumni would stop donations, and your university would decline. They turned the model into a monster as: “they had no direct way to quantify how a four-year process affected one single student, much less tens of millions of them. They couldn't measure learning, happiness, confidence, friendship, or other aspects of a student's four-year experience" (O'Neil, 2016) because college experience is an individual experience for self-fulfillment, can't be generalized and thereby, the model used proxies. The issue with proxies is that they are easier to manipulate compared with reality. For example, for a company that is hiring a social media manager, it's too time consuming to look over everyone's resumes and thereby may use a proxy, like who has the most Instagram followers. However, if this hiring parameter is known, every prospective candidate will work their way around it, buying followers and becoming more active in the platform. It forces everyone to aim for the same outcomes, entering a rat wheel. With universities, a lot of them which were the "safe" schools of many students, were thereby flooded with student applications, and because they knew what weighed in the algorithm, they started to decline a lot of applications to drive up their ranking, to become "more exclusive", according to the ranking. *Safe schools* would decline those brilliant prospects that would most likely get into their *target* school and not matriculate in their *safe* school, thus, keeping a lower acceptance rate. And perhaps, they are even losing those excellents students that might of prefered their *safe* school, the students that improve the educational experience for their classmates and even professors and now have to even derive financial aid to attract top students once again instead of giving it to those students that actually need it. The model evolved into a behemoth, swiftly becoming a national standard building a never-ending chain of self-destructive loops: "some administrators have gone to desperate lengths to drive up their rank. Baylor university paid the fee for admitted students to retake the SAT, hoping another try would boost their scores – and Baylor's ranking. Elite small schools, including Bucknell University in Pennsylvania and California's Claremont McKenna, sent false data to the *U.S. News* inflating the SAT scores of their incoming freshman. And Iona College, New York, acknowledged in 2011 that its employees had fudged numbers about nearly everything: test scores, acceptance and graduation rates, freshman retention, student-faculty ratio, and alumni giving" (O'Neil, 2016) and universities derived their focus from improving how the *U.S. News* algorithm saw them, instead of actually focusing on how to actually improve their education. This example of a model, although not as opaque as the models explained earlier, it is still biased and it harms the minorities, as Colin Diver describes: "Rankings create powerful incentives to manipulate data and distort institutional behavior for the sole or primary purpose of inflating one's score. Because the rankings depend heavily on unaudited, self-reported data, there is no way to ensure either the accuracy of the information or the reliability of the resulting rankings" (Diver, 2005) who used to be the dean of the Law School at the University of Pennsylvania and had to work under the shadow of *U.S. News* algorithm for ten years. Diver, who is now the president of Reeds College, one of the few institutions that declines to participate in the peer assessments and surveys that *U.S. News* uses for their rankings, explains further how: "ranking schemes undermine the institutional diversity that characterizes American higher education. The urge to improve one's ranking creates an irresistible pressure toward homogeneity, and schools that, like Reed, strive to be different are almost inevitably penalized" (Diver, 2005) and it is not just the uniformity rankins create but according to Reed's ethos, university education should produce personal satisfaction, the rankings reaffirm a perspective that education is purely instrumental to external rewards such as reputation. When Reed University opted not to provide any further data to *U.S. News,* they asked to be omitted from the ranking, instead, the magazine kept them in the rankings, using data from other sources and for the data that was missing, they decided to give them the lowest score possible, and how the editors came up with that value can only be speculated upon. Sadly, following Reed's lead seems like suicide; not compliying with the magazine results in a lower ranking but at the same time, the rankings create a distorsion of the educational value, and only five percent of universities don't comply with the magazine. Diver also explains the implications of answering the 656 peer questions and evaluations: "I'm asked to rank some 220 liberal arts schools nationwide into five tiers of quality. Contemplating the latter, I wonder how any human being could possess, in the words of the cover letter, "the broad experience and expertise needed to assess the academic quality" of more than a tiny handful of these institutions. Of course, I could check off "don't know" next to any institution, but if I did so honestly, I would end up ranking only the few schools with which Reed directly competes or about which I happen to know from personal experience. Most of what I may *think* I know about the others is based on badly outdated information, fragmentary impressions, or the relative place of a school in the rankings-validated and rankings-influenced pecking order" (Diver, 2005), meanwhile, Reed can focus on admitting those students that will promote an intellectual environment and not those who will plump the magazine's ranking. Furthermore, *U.S. News* puts a lot of weight on student retention, however, what this metric means regarding how an education is seems trivial and rewarding schools that keep their students and graduate them encourages them to focus on making them happy rather than on pushing them. Columbia's Math professor, Michael Thaddeus, published in February 2022 an article questioning Columbia's ranking in the *U.S. News* and how it has climbed from eighteenth place in 1988 to second place in 2022. Thaddeus concludes that: “several of the key figures supporting Columbia’s high ranking are inaccurate, dubious, or highly misleading” (Thaddeus, 2022) after analyzing the ranking model’s metrics and that: “discrepancies, sometimes quite large, and always in Columbia’s favor” (Thaddeus, 2022) between the data supplied for the ranking and the data supplied elsewhere. Now, a lot of businesses have erupted by which they help students get into universities. As each university has its own mini model of the *U.S. News* they use to accept applicants, new business like ThinkTank Learning emerge that understand what each university puts more emphasis on their model and helps customers tailor their profile to best fit the model, like how many community hours they need. ThinkTank Learning charges for their consultancy packages between $19,826 and $25,931 and parents are willing to pay this as they perceive it as a way to ensure success in their children’s acceptance in elite schools, Ma, founder of ThinkTank Learning explains on Bloomberg Business magazine how he is able to get students into Ivy Leagues . However, poor to middle-income families can’t afford this and can’t tailor their profiles to what the models are looking for, they are missing out on insider information and the outcome is an educational system that favors the privileged, keeping the social gap from disappearing.

While some universities duel with the *U.S. News* rankings, others benefited from it; thanks in part to the welcoming of Information and Communication Technologies (ICT). Those universities that derived their focus not on the “top” students, they targeted the poor, according to the article *96 percent of Google's revenue is advertising, who buys it?* by Meghan Kelly the University of Phoenix spent $50 million on Google ads, with the promise of rising higher socially and economically. In today’s world, everyone is categorized by models and targeted by their weaknesses or highest demans by ads (if you are short for money, suddenly the highest interest loans appear, if you are struggling in your sex live, suddenly advertisements or Viagra or Cialis appear). Sadly, the education that target those who are most desperate doesn't lead to a road of success, rather to a big pile of debt, and if you manipulate your "job placement" rates, "role students" acceptance rates and such, it's easier to drawn in with your adds the most vulnerable population and charging over priced.

7. PERSONALIZED ADVERTISEMENT

The Social dilemma, is a Netflix documentary that interviews early employees from Facebook, Apple, Twitter, Instagram, Pinterest and such. Now, they all explain through the screen why their ethical concerns made them leave and their concern with the industry at large. They explain how although these platforms had a positive effect, we are naive about the other side of the coin. The problems are vast: mental health and social media usage, ISIS and white supremacist able to inspire people online, fake news with consequences, elections getting hacked, how to handle an epidemic with fake news, and more. But what is the root behind all these problems?

A new agenda for technology, it shouldn't be just the tech industry that understands how this works, but rather that everyone knows. In this new era, a bunch of guys working at Silicon Valley have an impact on billions of people. Now, users might have ideas that they didn't intend to have because of how a designer decided how and which notifications you would wake up to. At the beginning of the technology revolution, Silicon Valley was based on selling hardware and software, now, it's based on selling user's information. We regard all these platforms (snapchat, Google, Facebook, Instagram, Reddit, Pinterest, Tiktok, etc) as free because we don't pay to use them when in fact, advertisers are paying for it, as the saying goes: "if you are not paying for the product, you are the product": we are the ones getting sold. All social platforms compete for your attention: their business model is to maintain people on the screen as long as possible and advertisers pay them to show us their ads, our attention is the product that they sell to the advertising companies. Furthermore, they get to choose what to display to you, gradually changing your own behavior: what you do, what you think and who you are. Harvard Business School Professor Shoshana Zuboff describes this new phenomenon in her book *The age of surveillance capitalism: The fight for a human future at the New Frontier of Power* where she explains how: "the stakes could not be higher: a global architecture of behavior modification threatens human nature in the 21st century just as industrial capitalism disfigured the natural world in the 20th" (Harvard Business School Review) because it explains on how international tech companies swayed us to surrender our privacy at the sake easiness: how data collected by these businesses is being used by others not only to predict but also to influence and modify our behavior; and how this has had catastrophic consequences for democracy and liberty. Zuboff's title refers to a new economic order and: "an expropriation of critical human rights best understood as a coup from above" (Zuboff, 2019) that is inserting itself from Silicon Valley into every economic sector with no opposition from laws nor society. This business model sells with a guarantee that a placed ad will be successful with complete certainty, and to do so they create predictions from vast amounts of data. Now, we not only have markets that sell on oil or other commodity futures, but human futures in economies of scale: making the technology companies the richest out of all. Today, everything you do is carefully monitored: from what you search to for how long you look at a picture, and they deduce personality traits such as if you are depressed or anxious and what you spent time looking at nights. All this data is introduced into algorithms to make better predictions about us and our decisions. What technology companies do is not sell our data but rather develop models that forecast our actions and whichever company creates the best model wins, because we will spend more time in their app/interface. If they notice our interest is decreasing they are able to trigger specific emotions that will keep us engaged and they know from all of our data what keeps us “scrolling”. One can summarize what technology companies do: getting us to engage in their platform as long as possible and creating a desire to log back again while monetizing this with the display of ads, and algorithms are in charge of getting the most revenue from each sesion you log in. In today's world, our socialization is first through the internet and there is a third person lucriating from it that at the same time is manipulating us. New generations are growing under persuasive technology, where they introduced psychology into technologies to manipulate our behavior: what it is called positive intermittent reinforcement in psychology can be seen in apps when you're sliding and decide to “refresh”, which is comparatively the same as the slot machines in casinos, which results into an unconscious habit. The fact that if you take your phone and a notification might display is by no means an accident, it's a design technique that exploits human vulnerability. Tech-companies are constantly doing A/B test on their platform, introducing small changes and seeing if it elicits responses from us and what response that is. We are lab rats but instead of trying to cure a disease, we’re just helping companies profit and we are clueless about it. Chamath Palihapitiya, who was Facebook VP of growth explains how they keep you in the loop: “we want to psychologically figure out how to manipulate you as fast as possible and then give you back that dopamine hit” (Lans, 2020) and although facebook started doing it, now all other platforms do it as well. Palihapitiya, who now feels tremendous guilt, devotes his time to encourage leaders to rein this beast in before it's too late to control it. Technology has shifted from being a tool to manipulating, seducing and demanding things from you to the point where it is an addiction. When something is a tool, you can leave it and only come for it when you need it, but this isn't the case with technology: we are not able to leave our phones and let them be, we are urging the moment a notifications pops up and it's hard to let “log off” and not keep using it. This is why we are called *users* and not “clients”, thus, there are only two industries that call us users, in drug consumption and on the Internet of things (IoT).

8. FACEBOOK

Any of the big giants in the tech industry, like Twitter or Facebook, might seem like a good resource for keeping up to date with today's news. According to Pew Research Report: “the share of Americans for whom Twitter and Facebook serve as a source of news is continuing to rise. [...]. The report also finds that users turn to each of these prominent social networks to fulfill different types of information needs” (Shearer, Barthel, Gottfried and Mitchell, 2015) which raises the question that if technology companies are tweaking and tailoring each algorithm for us, including the news we see, are they playing a part in the political system? From 2010 to 2012, researches at facebook described: “results from a randomized controlled trial of political mobilization messages delivered to 61 million Facebook users during the 2010 US congressional elections. The results show that the messages directly influenced political self-expression, information seeking and real-world voting behavior of millions of people. Furthermore, the messages not only influenced the users who received them but also the users' friends, and friends of friends. The effect of social transmission on real-world voting was greater than the direct effect of the messages themselves, and nearly all the transmission occurred between *close friends* who were more likely to have a face-to-face relationship" (Bond, Fariss, Jones, Kramer, Marlow, Settle and Fowler, 2012) concluded from seeing people post "I voted" and making it appear in 61 million people's feed, thus, Facebook was encouraging Americans to vote, plus, Facebook was inciting peer pressure by showing people's voting behavior. They were also analyzing how various types of updates impacted people's voting decisions. Facebook is just an example, but all big tech-companies have this much power and can steer an outcome: they are shaping the American government. Of course there is bias in newspapers or TV news channels, but when they cover a story; everyone sees it, it's not opaque. Most people believe that Facebook shares everything everyone says, according to a 2013 study by Karrie Karahalios: " 62% of people didn't know that their News Feeds were being filtered. When the algorithm was explained to one subject, she compared the revelation to the moment when Neo discovers the artificiality of The Matrix" (Luckerson, 2015), they were under the impression that the system instantaneously distributed anything they posted with their pals. Another research, this time on Google to demonstrate the strength and sturdiness of the effect of search engine manipulation (SEME) showed : " (*i*) biased search rankings can shift the voting preferences of undecided voters by 20% or more, (*ii*) the shift can be much higher in some demographic groups, and (*iii*) such rankings can be masked so that people show no awareness of the manipulation" (Robert Epstein and Ronald Hancock, 2015) and in part, this is because consumers have a high degree of faith in search engines: Pew

Research reported that: “73% of search engine users say that most or all the information they find as they use search engines is accurate and trustworthy” (Kristin Purcell, Joanna Brenner and Lee Raine, 2012). Everything we know about the big internet companies comes from the minuscule amount of research they share. Their algorithms are critical to their business.

Thanks to mass amounts of data and algorithms, microtargeting is implemented successfully, tailoring ads to different segments of the population. These more subdued campaigns are just as dishonest and less responsible. Microtargeting explains why 43 percent of americans believe the false pretense that President Obama is a Muslim (Bailey, 2014). Although political campaigns use mainly TV for advertisement, there is even a growing trend on television for profiling their viewers and adapting ads to them (Perlberg, 2014). As microtargeting grows, it's harder to know what media people have access outside your household and why they believe the things they do so passionately. This is called asymmetry of information, and it endangers cooperation between parties, which is exactly the point of a democratic government. Again, microtargeting is also opaque and unanswerable, which opens the pathway for politicians to adjust themselves to what each individual would appraise: they can portrait themselves as customized for each segment’s standards.

9. PERSONALITY TEST

Nowadays, most companies use personality tests for hiring, like the Kronos test, and people who get red-lighted almost never learn they were rejected from the position because of their personality tests. And yet, there are almost no legal proceedings to this method. The question is how automated systems evaluate us when we apply for jobs and what criteria they use to do so. So far, we've seen algorithms impact negatively on college admissions and the criminal justice system and how this impacts each racial and ethnic group. What they all have in common is that they need a job, whether it is in a supermarket or a bank. Naturally, these hiring tools are unable to incorporate information regarding a candidate's real performance in the workplace. That is the future, and so unknowable. As a result, they, like many other Big Data tools, rely on proxies. And, as we have described, proxies are invariably imprecise and frequently unjust. In 1971, the Supreme Court ruled in *Griggs v. Duke Power Company* (Griggs v Duke Power, 2018) that hiring based upon intelligence tests were discriminatory and thus illegal. However, this only led to the replacement of intelligence tests to personality tests. Frank Schmidt, a business professor, who has evaluated more than a century's worth of productivity data in order to determine the predictive value of various selection procedures (The problem with using personality tests for hiring, 2017) showed on his research that personality tests are poor predictors of job performance: they displayed far lower than reference checks and were one third as predictive as cognitive exams. This program is used to rule out the most people possible at the cheapest cost; not to find the perfect candidate. And again, these personality tests are opaque because when they ask their questions, like: “would you describe yourself as distinct or organized?” candidates have to choose an option without knowing how a machine will interpret those, and some conclusions might result unfavorable. If you are red-lighted by Kroger and later on become an outstanding employee somewhere else, no one will go back and look into why Kroger failed with your predictions. Again, this model is different from those of like baseball teams, because they look at every potential individual as they each could be worth millions, and if they end up not selecting a player but then he outperforms in another team, they have to look back at their analytic engine: the difference is clear constant feedback. However, most companies are managing masses, not individuals and they prefer to cut expenses by replacing human resource professionals with machines, because at the end of the day they are completing the job, even if they end up rejecting future prodigies. However, although companies might be fine with this status quo, again, there are real people affected by these systems. Moreover, it is unknown to most job applicants that: "72% of resumes are never seen by human eyes" (Abdel-Halim, 2012) because of these systems that filter resumes, and in addition, more and more companies: "are beginning to embrace a concept called *workforce science* that promises to make performance reviews and judging résumés far more data-driven. One of the best-known attempts to hire and fire by the numbers is at Google, whose HR department, called *People Operations*, has turned hiring into a kind of engineering project, using computer models to determine how many times each candidate should be interviewed, how larges raises should be, and nearly every other personnel decision" (Leber, 2020), bitterly, algorithmic programs are carving more amore their way into the decision making matrix for the workforce. Before these programs were in place, of course hiring managers had biases, however, these biases differed from company to company, and people affected by one biased could still have an opportunity in another company, while now, if you are red-lighted by a system, you while be red-lighted by every company that uses personality tests. Furthermore, a person with mental health problems will be discriminated against by these personality tests, which is a violation of the American Disability Act, which is intended to avert these people from not having a normal life, thus, integrating them into society, not isolating them.

10. GETTING CREDIT

Before, if you wanted to land a loan, you would put on your best clothing and visit your local banker, while your banker also took into account what he knew about you in your community: your lifestyle, if you went to church, your family's history and your reputation around town. Thus, the banker had his own preconceived judgements about you, and minorities suffered bias as well this way. It wasn't long before algorithms found their way into this sector: FICO was a model used to determine the likelihood of an individual defaulting on a loan. This FICO score was calculated using a procedure that considered solely the borrower's financial situation. The score was color blind, which benefited the banking industry significantly, since it forecasted risk far more correctly while also introducing millions of new customers. The FICO: " is currently the standard metric in circulation for evaluating consumers in the U.S. market for consumer credit" (Poon, 2007). The FICO score is now used by a large number of credit bureaus, which each add unique data sources to the FICO model in order to generate their own scores. These scores exhibit a variety of desirable characteristics. To begin, they have a well-defined feedback loop. Credit agencies can track which borrowers default on their loans and compare those statistics to the borrowers' credit scores. If it appears that borrowers with high credit scores are defaulting on loans more often than the model predicts, FICO and the credit reporting agencies can adjust the algorithms to improve their accuracy. Additionally, credit ratings are fairly transparent. FICO's website includes straightforward directions on how to boost your score (FICO's website).

Additionally, the industry of credit scores is controlled. If you have a query about your score, you have the legal right to get a copy of your credit report (Hebert, Hernandez, Perkins and Puig, 2022), which contains all of the data used to calculate the score. Though the procedure can be lengthy, if errors are discovered, they can be corrected. Pseudoscientific methods, on the other hand, aim to anticipate our creditworthiness by assigning us an e-score. These rarely seen numbers unlock doors for some of us while slamming them shut for others. In comparison to FICO ratings, e-scores are arbitrary, unaccountable, uncontrolled, and frequently unfair. Credit scores, which are based on personal credit reports, have existed for years. Additionally, direct marketing organizations have historically classified consumers according to their socioeconomic position. However, e-scores go a step farther, as Natasha singer explains, when you make a phone call, the e-score rapidly enlist callers in an order and: “they can determine whether a customer is routed promptly to an attentive service agent or relegated to an overflow call center” (Singer, 2012) depending on their profitability; if your score shows you are more likely to buy a product or service, an agent will soon answer your, however, if your scores shows your not as likely to make a purchase, you’ll go to the end of the line or just dealt with by automatic machines. Regulatory agencies and consumer activists are concerned that these rankings would eventually disadvantage some customers, particularly those experiencing financial hardship. In effect, they argue, the ratings could create a new default class of people: those who are unknowingly overlooked by businesses online. Financial firms, in particular, may ignore individuals with low credit ratings, limiting their access to mortgages, credit cards, and insurance. Another example, by credit cards, is Capital One, who uses such calculations: “to instantly decide which credit cards to show first-time visitors to its website" (Steel and Angwin, 2013) and as the article mentions, the director behind these calculations says they: "never know nothing about a person" (Steel and Angwin, 2013), they can make assumptions based on what type of car you search for or from which location, also known as *geo-tag*, are you "surfing" the web and comparing them with other users: a person searching for a Range Rover is most likely wealthier than the one searching for a Toyota and they draw this type of conclusions. Consider the negative spiral created by e-scores. There's a good likelihood that the borrower from the rough area will receive a low score from the e-scoring system. There are many people who have defaulted there. As a result, the credit card offer that appears on their screen will be aimed at a more risky demographic. For those who are already suffering, this means fewer lending options and higher interest rates. And because it's illegal to use credit scores for marketing, they use these substitute e-scores. It makes sense that our credit scores are hidden as they contain a lot of sensitive information; nevertheless, as a result, firms end up digging into mostly uncontrolled pools of data, such as geo-tags, in order to build a parallel data marketplace In the pre-FICO banking era, bankers drew conclusions from individual, like race and church habits and segmented people into "buckets": some people are trustworthy and others are not. When Fair and Isaac created the FICO score, those previous proxies were thrown out as it focused on the individual and not in what bucket a person would fit in. e-score, on the other hand, takes us backwards in history. They examine the individual using plenty of proxies. They perform thousands of "people trustworthy" estimates in instants. E-scores answer the question "how have people similar to you acted in the past?" whereas FICO answers: " how have you acted in the past?" and the distinction between these two inquiries is enormous. Of course there are some traits similar people tend to do and that are correct to assume, wealthier people tend to buy a Land Rover, however, when someone is done wrong, the wrongdoings don't go into the algorithm to make it better. Now, if someone looses their job, they can't paid their bills, but they might not be able to find a new job because they've developed a bad credit score, according to the Society for Human Resource Management: "nearly half of employers use credit checks when making a hiring decision, according to a 2012 survey " (Rivlin, 2013), which creates a job discrimination as most minorities tend to have low scores: its becoming harder and harder for latino,black and poor people to get hire making it harder for minorities to get out of poverty. Employers may argue that a good credit score is a sign of trustworthiness in a employer, however, there's a lot of trustworthy and with good morale people that might have a bad score: firing people due to recessions or to cut cost or having an injury or illness at a time that you didn't have insurance. These are circumstantial events that shouldn't define a person, and when a rough patch happens, those living paycheck to paycheck, like minorities, can't keep a good credit, unlike those who are able to save for hard times. Good credit is used as a proxy, however, good credit is more related to wealth than trustworthiness, and sadly, wealth is also related to race. According to a CNN article: "when it comes to wealth, the difference is staggering. Whites have roughly 10 times the wealth of blacks and Hispanics. Over the past 25 years, the wealth gap between blacks and whites has nearly tripled, according to research by Brandeis University" (Luhby, 2013) and they also show data from the bureau of labor statistics that reveals that white people have the lowest unemployment rate (4.4%) and the highest home ownership (71.9%) and highest median household wealth ($134,230) while black people have the highest unemployment rate (9.2%) and the lowest home ownership (42.2%) and median household wealth ($11,030).

In the IoT, a lot of work is automated, and even the best algorithms can present errors that may hunt you for the rest of your life, unless you are able to first, detect them and second, ability and time to fix them: “inaccuracies often show up in consumers’ credit reports, and these errors have real consequences, like increasing borrowing costs or barring people from financing a home or renting an apartment. And once an error is found, getting it fixed can take months of exasperating work”(Morgenson, 2014). This is the case for Patricia Armour, from Mississippi, who spent over two years trying to fix her Experian credit report: “a second mortgage that had been discharged when she filed for bankruptcy in 2007 popped up as an unpaid debt of around $40,000 in 2011, she said. Even though she repeatedly supplied proof of the discharge to Experian, she said, it refused to fix the error” (Morgenson, 2014) and not till after she contacted the Mississippi’s attorney general’s office was she able to correct the report. Under the Fair Credit Reporting Act: “these agencies are supposed to have procedures assuring *maximum possible accuracy* of consumers’ information. The law allows consumers to check the reports for errors and requires credit bureaus to investigate consumers’ error claims” (Morgenson, 2014), however, as the Mississippi attorney general concludes, Experian has engaged in an unbroken habit and practice of breaking state and federal law for decades, a pattern they seem to refuse to fix, as they are willing to pay for all their settlements and law procedures, but reluct to take the required measure to avoid this: it is faster and cheaper to wait for someone to realize they have been wrong, try suing them with the costs and time this implies than to fixing the problem: “experian generated $4.8 billion in revenue for the year ended March 2014, and its after-tax profit of $747 million in the period was more than twice its 2013 figure” (Morgenson, 2014), the price the pay in settlements is little compared with the profits they make, which makes it understandable why they are so resistant to change, that's their business model. In the work at the general attorney's office they: "spent more than a year interviewing former employees and reviewing complaints about Experian from state residents. Investigators found that the company routinely mixed up reports of consumers who have the same name, allowed erroneous information to be included on credit reports and would not correct the errors that consumers had identified. The company also failed to investigate disputed data as required, the complaint said, and accepted creditors' findings about the disputed information even if it was contradicted by canceled checks or other proof" (Morgenson, 2014), and as result, consumer are damaged the most: from not getting a credit to paying higher interest rates. Gretchen Morgenson, who goes on explaining the issues with flawed credits in the New York Times' article *Held Captive by Flawed Credit Reports,* concludes that: " there is no doubt that erroneous information on credit reports remains an enormous problem. Last year, the Federal Trade Commission found that 5 percent of consumers — or an estimated 10 million people — had an error on one of their credit reports that could have resulted in higher borrowing costs" (Morgenson, 2014), and from 2000 to 2014, 18 actions were filed against reporting bureaus she says. At the end of the day, consumers here are not the reporting agencies' primary source of revenue; credit providers are. The information is paid for each time a customer applies for a loan, mortgage or lease. Due to the fact that consumers are not their genuine clients, bureaus have little motivation to treat them effectively. As an Experian employee explains: "they were pressured to meet *production* quotas and given no more than five minutes to handle each consumer call. These employees also described internal competitions for speedy call-handling, bonuses for meeting quotas and probation for those with low production numbers" (Morgenson, 2014), yet another impediment to improved treatment. Although it might take time, at least with credit reports, there is a regulated side.

Even more perilous is the uncontrolled side of the Big Data industry. Numerous corporations purchase data from shops, advertising, smartphone app developers, and social network operators in order to compile a buffet of facts about every consumer , such as a customer with cancer. Additionally, these corporations collect any publicly available government data, including criminal records. All of this information is compiled into a customer profile that they sell. Certain data brokers are more trustworthy than others. However, any process that attempts to profile millions of citizens using lots of different sources is bound to make several errors. For example, Helen Stokes, 63 year old living in Philadelphia could no longer afford her big home: "but her search for a subsidized apartment was haunted by an arrest record that legally didn't exist" (Pallazzolo and Fields, 2015), her arrest were due to fights with her now ex-husband and none resulted in a conviction and she asked to get her records expunged. Ms. Stoke is a case of many that although after asking to get her files out of the government database, still prevail in the private sector: "a company called RealPage Inc. that provides background screening for landlords continued to report her arrest records six months later, leading two senior living centers to reject her applications" (Pallazzolo and Fields, 2015) and there is no data available as to how often this happens as per definition, spunged records no longer exist. RealPage and such types of companies make money out of creating and selling people's reports, and Ms. Stoke is not a customer but the product, and she didn't get her the RealPage to clear the information till after she sued, which again, takes a lot of money and time, and yet, we can't be sure that the RealPage isn't still selling her profile with that information.

Automation still misses out on a lot of things, and they store errors in our consumer's profiles. Google, one of the best at artificial intelligence and machine learning, launched this feature called photo-tagging, which does automatic photo tagging, and it tagged two black people as *gorillas*: "the gorilla tags turned up in the search feature of the Google Photos app, which the company released a few weeks ago. When users start a search, Google suggests categories developed from machine learning, the science of training computers to perform human tasks such as labeling. The company has removed the gorilla categories, so those suggestions will no longer appear" (Barr, 2015), which was probably a result of the new feature probably not thoroughly tested prior to release, leading to a faulty machine learning that confused *Homo sapiens* with gorillas from scanning through thousands of primate images and constructed its own characteristics. Accidental errors are instructive, as long as the system obtains feedback about the error, which it did in this case. However, unfairness continues. When automated algorithms dig over our data in order to generate an e-score for us, they inadvertently project the past further into the future. As with sentencing models for recidivism and loan algorithms, the poor are assumed to remain poor in perpetuity and are treated as such. We cannot rely on automated systems to resolve the problem. Despite their astounding capabilities, machines cannot yet make adjustments for justice on their own. Sifting through data and determining what is just is both unfamiliar and incredibly complex to them. Only human beings are capable of enforcing that limitation. Today the world is governed by automated systems, which were created to avoid human bias and enforce fairness, however, this development, although it seems it was done, hiddes further prejudice and errors. Only human beings can provide common sense to this new frame.

If, on the other hand, we leave this matter to the market, which values speed, productivity, and profit while accepting some significant errors, interfering people will be taught to stay away from the technology, which will become an issue as strong entrants flood in. Facebook, has a new patent for creditworthiness based on users' Facebook friends. The obtained pattern, which lets Facebook analyze user's friends, has a list of things they do with this analyzed data, one of them at the end of the list being: "when an individual applies for a loan, the lender examines the credit ratings of members of the individual's social network who are connected to the individual […]. If the average credit rating of these members is at least a minimum credit score, the lender continues to process the loan application. Otherwise, the loan application is rejected" (Meyer, 2015), which leads to our friends shaping our credit profile, and leading to the question if this should even be legal.

11. FINAL CONCLUSIONS

In this research paper it has been discussed how algorithms and mathematical models play a big role in colleges, the courts, working space, voting and even credit loans. With the promise of efficiency and justice, they corrupted higher education, accelerated mass incarceration, targeted the poor at practically every opportunity, and undermined democracy. Being impoverished in an algorithm-driven society is becoming increasingly risky and costly. This exploitation of the impoverished also isolates the affluent classes of society from one another in terms of marketing. For most of them, it can seem as though the world is becoming smarter and faster: obtaining movie recommendations or avoiding "bad" shopping zones. The discreet and personal nature of this targeting obscures the fact that the precise same models are ruining lives just a few blocks away. Mathematical models operate in the dark and they are opaque: slicing into many and concealing the harm they cause to others. Damage, opacity, and scaling are all features in algorithms. They have the appearance of being a “black box” because they are generally proprietary or otherwise protected from outside scrutiny. Because they influence so many people, there's a higher likelihood that they'll get it incorrect for some of those. Racism or other biases may be encoded into an algorithm to make it easier for unscrupulous firms to target specific groups of vulnerable individuals, or they may even cause a worldwide financial collapse. Dismantling obfuscated algorithms may not always result in such a clear benefit. While more justice and fairness would undoubtedly benefit society as a whole, individual businesses are not in a position to reap the benefits. In reality, these algorithms appear to be quite successful for the vast majority of them. They're the foundation of whole company concepts. And those who profit from a software program that successfully targets those desperate enough to pay higher rates believe it is functioning well. The victims, on the other hand, have a different perspective. However, the vast majority of them are poor, including hourly employees and the jobless, as well as those with bad credit ratings. Prisoners have no control. These algorithmic casualties are almost silent in our culture, where money buys power. Politically, the majority are alienated. Indeed, the poor are frequently blamed for their poverty, inadequate schools, and the violence that plagues their communities. As a result, few politicians bother with anti-poverty policies. Poverty is viewed as a sickness, and attempts are being made to isolate it and prevent it from spreading to the middle class. We must consider how we allocate blame in modern society, as well as how models intensify this loop. However, the poor are far from the only ones who have suffered as a result of Big Data. We've previously seen how nefarious models may eliminate qualified job candidates. The middle class is hurt by big data as well. And in order to appease the merciless algorithms that dominate university admissions and poison higher education, they scamper about as furiously as the rest.

We have repeatedly demonstrated that mathematical models can skim through data to identify individuals who are likely to face significant obstacles, whether related to crime, poverty, or education. It is up to society to decide whether to refuse and penalize them or to reach out and provide them with the resources they require.

It's also worth noting that we're still in the early stages of this industry. Of course, they begin by going after the poor, who are the easiest prey to take advantage of due to their lack of resources and general desperation. But these mathematical models that generate fantastic profit margins are not going to remain in the lower tiers for long. In reality, this isn't how markets operate. They'll continue to grow and change as they seek out new possibilities. Mathematical models, in summary, are aimed at everyone. They'll keep multiplying and demonstrating injustice unless we do something about it. Even in the age of Big Data, you could claim that human cruelty is no worse than it has ever been. Many people believe that even the worst mathematical models aren't all that horrible. However, despite its flaws, human decision-making has one major virtue. It has the capacity to develop. Humans, like all living things, are always evolving and learning new things. Automated systems, on the other hand, are immobile until engineers get their hands dirty and intervene. Because if a Big Data college admission model had existed in the early 1960s, the number of women enrolled in higher education would have remained low due to the model's reliance on successful male role models. The history is codified thanks to big data operations. Neither they, nor anybody else, can foresee the future. That demands moral imagination, which can only be provided by people. We must actively include better morals into our algorithms in order to create Big data models that follow our moral lead. That may necessitate sacrificing profit for the sake of justice.

In a way, we're in the midst of a new industrial revolution in our civilization. We can also take some lessons from the prior case. Things like coal-fired power and heat made life better for many people. However, there was a terrible side to this advancement. Powered by terribly mistreated employees, many of whom were kids. Coal mines were particularly hazardous before health and safety rules were implemented. Customers' rates went risen when railways, energy firms, and utilities were ruled by a monopolist. Without a doubt, the free market was unable to rein in its overproduction. There were no safety measures or child labor laws until individuals began to protest about the conditions in the factories, and hence, the government stepped in.

To begin regulating these mathematical models I believe the approach should begin with the modelers themselves. Data scientists, like doctors, should vow against misuse and be wary of model assumptions. Solid morals and consciousness, on the other hand, restrain only the most committed. We must go beyond creating industry standards in our data industry to remove faulty algorithms. Our laws must also evolve. And in order to do so, we must reconsider our success criteria. Today, a model's performance is frequently assessed in terms of profit, effectiveness, or debt levels. Human values should be imposed on these systems even if it means sacrificing effectiveness. In addition, we must assess the impact of these models and carry out algorithmic inspections. Despite being in the age of Big Data, it is still an issue that can only be solved by humans.

Data will not vanish. Neither are computers or math. Predictive models are becoming an increasingly important tool for running our organizations, allocating resources, and managing our lives. However, as demonstrated in this work, these models are built not only from data but also from decisions we take about which data to pay close attention to and which to ignore. These decisions are about more than just operations, business, and productivity. They are moral in their core. We abdicate responsibility if we distance ourselves from them and accept mathematical models as an impartial and unavoidable force. As a consequence, which we've seen, algorithms treat humans like machine components at work, blackballing employees and feasting on injustices. We must get together to regulate these mathematical models, taming and disarming them.

12. BIBLIOGRAPHY

Abdel-Halim, M. (2012, May 27). *12 ways to optimize your resume for Applicant Tracking Systems*. Mashable. Retrieved April 20, 2022, from https://mashable.com/archive/resume-tracking-systems

Bailey, S. P. (2015, September 14). *A startling number of Americans still believe president Obama is a Muslim*. The Washington Post. Retrieved April 22, 2022, from https://www.washingtonpost.com/news/acts-of-faith/wp/2015/09/14/a-startling-number-of-americans-still-believe-president-obama-is-a-muslim/

Barr, A. (2015, July 1). *Google mistakenly tags black people as 'gorillas,' showing limits of algorithms*. The Wall Street Journal. Retrieved April 30, 2022, from https://www.wsj.com/articles/BL-DGB-42522

Barry-Jester, A. M. (2015, August 4). *Should prison sentences be based on crimes that haven't been committed yet?* FiveThirtyEight. Retrieved March 6, 2022, from https://fivethirtyeight.com/features/prison-reform-risk-assessment/

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012, September 12). *A 61-million-person experiment in social influence and political mobilization*. Nature News. Retrieved April 21, 2022, from https://www.nature.com/articles/nature11421

Calo, R. (2014, August). *Digital Market manipulation - gwlr.org*. Retrieved April 24, 2022, from https://www.gwlr.org/wp-content/uploads/2014/10/Calo_82_41.pdf

Diver, C. (2005, November 1). *Is there life after rankings?* The Atlantic. Retrieved March 20, 2022, from https://www.theatlantic.com/magazine/archive/2005/11/is-there-life-after-rankings/304308/

Epstein, R., & Hancock, R. (2015, August 4). *The search engine manipulation effect (SEME) and ... - PNAS*. Retrieved April 22, 2022, from https://www.pnas.org/doi/10.1073/pnas.1419828112

FICO's website. *The #1 analytic decisioning platform to optimize consumer interactions across all customer decisions*. FICO. Retrieved April 25, 2022, from https://www.fico.com/

Griggs v Duke Power. NAACP Legal Defense and Educational Fund. (2018, November 15). Retrieved April 19, 2022, from https://www.naacpldf.org/case-issue/griggs-v-duke-power-co/

Hebert, A., Hernandez, A., Perkins, R., & Puig, A. (2022, March 29). *Free credit reports*. Consumer Advice. Retrieved April 25, 2022, from https://consumer.ftc.gov/articles/free-credit-reports

Kelly, M. (2012, January 30). *96 percent of Google's revenue is advertising, who buys it? (infographic)*. VentureBeat. Retrieved March 24, 2022, from https://venturebeat.com/2012/01/29/google-advertising/

Lans, P. byS. van der. (2020, January 2). *Transcript: Facebook's dopamine-driven feedback loops*. Sebastiaan van der Lans. Retrieved April 18, 2022, from https://sebastiaans.blog/facebook-dopamine-driven-feedback-loops/

Leber, J. (2020, April 2). *The Machine-Readable Workforce: Companies are analyzing more data to guide how they hire, recruit and promote their employees*. MIT Technology Review. Retrieved April 20, 2022, from https://www.technologyreview.com/2013/05/30/178270/seven-must-read-stories-week-ending-may-31-2013/

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2020). Artificial intelligence in business: State of the art and future research agenda. *Journal of business research*.

Luckerson, V. (2015, July 9). *Here's how Facebook's News Feed actually works*. Time. Retrieved April 22, 2022, from https://time.com/collection-post/3950525/facebook-news-feed-algorithm/

Luhby, T. (2013, February 27). *Wealth inequality between blacks and whites worsens*. CNN. Retrieved April 29, 2022, from https://www.cnn.com/2013/02/27/us/wealth-inequality-between-blacks-and-whites-worsens/index.html

Martin, K. E. (2019). Designing ethical algorithms. *MIS Quarterly Executive June*.

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835-850.

Martin, K. E. (2015). Ethical issues in the big data industry. *MIS Quarterly Executive*, *14*, 2.

Meyer, R. (2015, September 25). *Could a bank deny your loan based on your facebook friends?* The Atlantic. Retrieved May 1, 2022, from https://www.theatlantic.com/technology/archive/2015/09/facebooks-new-patent-and-digital-redlining/407287/

Michael Lewis, Moneyball: The Art of Winning an Unfair Game (New York: W.W. Norton, 2003).

Morgenson, G. (2014, June 21). *Held captive by flawed credit reports*. The New York Times. Retrieved April 30, 2022, from https://www.nytimes.com/2014/06/22/business/held-captive-by-flawed-credit-reports.html

Palazzolo, J., & Fields, G. (2015, May 7). *Fight grows to stop expunged criminal records living on in background checks*. The Wall Street Journal. Retrieved April 30, 2022, from https://www.wsj.com/articles/fight-grows-to-stop-expunged-criminal-records-living-on-in-background-checks-1430991002

Perlberg, S. (2014, November 20). *Targeted ads? TV can do that now too*. The Wall Street Journal. Retrieved April 22, 2022, from https://www.wsj.com/articles/targeted-ads-tv-can-do-that-now-too-1416506504

Poon, M. (2017). *Scorecards as devices for consumer credit: the case of Fair, Isaac & Company Incorporated*. Retrieved April 25, 2022, from https://journals.sagepub.com/doi/full/10.1111/j.1467-954X.2007.00740.x

Purcell, K., Brenner, J., & Rainie, L. (2012). *Search engine use 2012kRISTIN*. Pew Research Center: Internet, Science & Tech. Retrieved April 22, 2022, from https://www.pewresearch.org/internet/2012/03/09/search-engine-use-2012-2/

Humerick, J. D. (2020). Reprogramming Fairness: Affirmative Action in Algorithmic Criminal Sentencing. HRLR Online, 4(2), 213-245.

Rivlin, G. (2013, May 11). *The Long Shadow of bad credit in a job search*. The New York Times. Retrieved April 29, 2022, from https://www.nytimes.com/2013/05/12/business/employers-pull-applicants-credit-reports.html

Steel, E., & Angwin, J. (2013, January 3). *Anonymity in name only - tracking technology on the web*. The Wall Street Journal. Retrieved April 28, 2022, from https://www.wsj.com/articles/SB10001424052748703294904575385532109190198

Shearer, E., Barthel, M., Gottfried, J., & Mitchell, A. (2015, July 14). *The evolving role of news on Twitter and Facebook*. Pew Research Center's Journalism Project. Retrieved April 21, 2022, from https://www.pewresearch.org/journalism/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook/

Singer, N. (2012, August 18). *Secret e-scores chart consumers' buying power*. The New York Times. Retrieved April 28, 2022, from https://www.nytimes.com/2012/08/19/business/electronic-scores-rank-consumers-by-potential-value.html

Thaddeus, M. (2022). *An Investigation of the Facts Behind Columbia's U.S. News Ranking*. Columbia and U.S. news. Retrieved March 20, 2022, from https://www.math.columbia.edu/~thaddeus/ranking/investigation.html

The age of surveillance capitalism: The fight for a human future at the New Frontier of Power. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power - Book - Faculty & Research - Harvard Business School. Retrieved April 18, 2022, from https://www.hbs.edu/faculty/Pages/item.aspx?num=56791

The problem with using personality tests for hiring. Harvard Business Review. (2017, December 6). Retrieved April 19, 2022, from https://hbr.org/2014/08/the-problem-with-using-personality-tests-for-hiring

Waldman, P. (2014, September 4). *How to Get into an Ivy LeagueCollege*. Bloomberg.com. Retrieved March 21, 2022, from https://www.bloomberg.com/news/articles/2014-09-04/how-to-get-into-an-ivy-league-college-guaranteed

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for the future at the New Frontier of Power*. Profile Books.