Spring 5-21-2022

Data Ethics: An Investigation of Data, Algorithms, and Practice

Gabrialla S. cockerell
Seattle Pacific University

Data Ethics:

An Investigation of Data, Algorithms, and Practice

Gabrialla Cockerell

Seattle Pacific University

May 6, 2022

ABSTRACT

This paper encompasses an examination of defective data collection, algorithms, and practices that continue to be cycled through society under the illusion that all information is processed uniformly, and technological innovation consistently parallels societal betterment. However, vulnerable communities, typically the impoverished and racially discriminated, get ensnared in these harmful cycles due to their disadvantages. Their hindrances are reflected in their information due to the interconnectedness of data, such as race being highly correlated to wealth, education, and location. However, their information continues to be analyzed with the same measures as populations who are not significantly affected by racial bias. Not only can the data itself be manipulated by collection methods, but faulty algorithm design and poor practices and implementation can intensify the damage. Such as lack of: regulation, racial literacy, peer review, expiration of data, privacy retention, and more. There is a denial of privacy for those who are vulnerable such that they cannot afford privacy and require more attention in general, which leads to oversurveillance. Additionally, there is an imbalance of value between the analytic information of subjects and the actual humans creating such data, so that the value of bodies is diminishing in quantitative arenas.

INTRODUCTION

Data today is in high demand, blooming with potential, and produced constantly. According to Forbes in 2018, 90 percent of the total data available to the world was produced in the previous two years[1], and this statistic continues to increase. So, as data increases at immense

[1] Marr, B., 2018. "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read." *Forbes*. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=10ce76c960ba.

rates, as well as the potential application and profit, so does the potential harm. The mathematics of data science is unfortunately not only tied to the struggles of the world as potential solutions, but also "fueling many of them."[2] Data science is an extremely widespread apparatus that has become increasingly complex given the sensitivity of personal information. When data is collected, analyzed, or implemented poorly, it can have resounding negative impact on a vulnerable subject's life because it weakens them further. We will expose a multitude of faulty attributes of data science today, and also gauge possible routes of mitigating the harm. Data models are predominantly designed with the intention for improvement for all, but it can be extremely easy to negatively impact the people who are at the mercy of algorithms.

CASE STUDIES

An efficient exploration of the ethics of data, algorithms, and practices requires an introduction to our case studies for context. Initially, we will explore risk assessments of inmates, predictive crime models of neighborhoods, and teacher evaluation scores. Eventually we will also examine financial risk algorithms for credit cards like AMEX and Capital One, and general hiring algorithms.

Criminal Risk Assessments and Predictive Crime Algorithms

The US criminal justice system commonly utilizes predictive crime algorithms and criminal risk assessments in attempt to prevent future offenses by processing the data of previous offenders and potentially risky areas and communities.

[2] O'Neil, C., 2017. *Weapons of Math Destruction*. Harlow: Penguin Books, p.2.

The LSI-R is a common inmate questionnaire[3] designed to assess the level of risk an inmate can impose on their community, as well as helping the penal system determine the next procedure of their sentence and rehabilitation to eventually reenter their community. These assessments are meant to be unbiased to any particular demographic, weighing each inmate equally based on a series of personal questions, such as frequency of police interactions, first involvement with police enforcement, domestic address, education, and financial history[4]. It is unethical to use race as a deciding factor in this process, yet we must consider all the byproducts of generational disparities between races that will inherently impact these models.

People of color are unequally incarcerated in our penal system. For example, "sentences imposed on black men in the federal system are nearly 20 percent longer than those for whites convicted of similar crimes."[5] Externally, it appears that the fair solution is the employment of these supposedly impartial models, but even the inquiry of home location can begin to racially divide offenders.

In America, our history has seen the development of generational wealth disparities, widely due to residential segregation. Despite the abolition of Black slaves, the immediate development of Jim Crows laws significantly barred their abilities to gain wealth as they lacked access to well-maintained institutions and sufficient salaries for their work[6]. Additionally, Black Americans were declared uninsurable by famous American statistician Frederick Hoffman in

---

[3] "The LSI-R is a standardized actuarial instrument that contains 54 items and produces a summary risk score that can be categorized into five risk levels… ranges have been designated that indicate an individual's risk category." *Federal Probation*, 2007. "The Predictive Validity of the LSI-R on a Sample of Offenders Drawn from the Records of the Iowa Department of Corrections Data Management System." 71(3), p.2. https://www.uscourts.gov.

[4] Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Oxford, England: Polity.

[5] O'Neill, *Weapons of Math Destruction*, 24.

[6] Hansan, J., 2011. "Jim Crow Laws and Racial Segregation." *Social Welfare History Project.* https://socialwelfare.library.vcu.edu/eras/civil-war-reconstruction/jim-crow-laws-andracial-segregation/.

1896, since they statistically had lower life expectancies at this point[7]. However, Dr. Hoffman failed to acknowledge that their preexisting significant poverty levels and Jim Crow laws—that denied financial opportunities—had led to less available medical service and poor living conditions, which directly affected their life expectancy. When their housing opportunities were restrained from better neighborhoods both because of their skin color and lack of mortgage loans, the type of areas and properties people of color qualified for were extremely restricted, which resulted in redlining. Since they were forced to buy or rent in lower income neighborhoods, they appreciated less on investment properties than White borrowers, enacting a dangerous feedback loop of poverty.

Thus, even when the Fair Housing Act of 1968 was passed to prohibit racial discrimination, generations of wealth disparities already made them often unqualified for credit opportunities, and they continued to locate in the same areas typically. Thus, this trend continued to bleed into the following generations, and they have appreciated less wealth on average, so that zip codes are inherently tied to racial divisions today[8].

In addition, the oversurveillance of Black communities can be traced back over a century ago as well, with watching of slaves and freed Black citizens to prosecute 'black misconduct'. This trend even bled into the civil rights movement of the 1960's, with the intentional observation of "race agitators" like Martin Luther King Jr., Malcolm X, and the Black Panthers.[9] Today, predictive crime algorithms show that nonwhite citizens of poor neighborhoods are more likely to commit crime, but conveniently do not consider their poor public-school opportunities

---

[7] O'Neill, *Weapons of Math Destruction*, 161.
[8] Franco, J. and Mitchell, B., 2018. "HOLC "redlining" maps: The persistent structure of segregation and economic inequality." *NCRC*. https://ncrc.org/holc/.
[9] Dennis, A. and Martin, J., 2020. "Mass Surveillance and Black Legal History." *American Constitution Society*. https://www.acslaw.org/expertforum/mass-surveillance-and-black-legal-history/.

can drive up crime rates due to overall neighborhood poverty, as "from a rational choice perspective, a causal link exists between poverty and property crimes because generating wealth reduces the benefit-to-cost ratio of committing crimes."[10] And this kind of predictive data model implements neighborhoods' existing crime rates to assess which areas are more at risk.

Crime rates are already ramped up in impoverished neighborhoods because increased frequency of police interactions in general will inevitably lead to increased crime rates in their communities as officers are hypersensitive to these areas. Young black men are more likely to have earlier and more frequent run-ins with the police than their Caucasian counterparts because of this imbalance of surveillance in their neighborhoods, not to mention the increased likelihood of being charged for the same activity[11]. This reality further develops racial stereotypes and increases policing in lower income areas even more. So, when questionnaires demand information like an inmate's first engagement with law enforcement and frequency of interactions, they can be punished for coming from a neighborhood that is predisposed to increased surveillance instead judged according to their individual actions. All the while, "prisoners are completely unaware of their questionnaire scores,"[12] which shields them from the internal processing of their personal information.

Surveillance data has even become a tool to the public, which further reinforces institutional racism. The average citizen can be weaponized with mobile applications like SketchFactor, developed to allow users to report "sketchy" areas so other people can avoid them.

---

[10] Bender, A., 2021. "How Poverty Drives Violent Crime." *OK Justice Reform.* https://okjusticereform.org/2021/12/how-poverty-drives-violent-crime/.
Mahnken, K., 2019. "Failing Schools: Home to Underachieving Students, Disillusioned Teachers and — According to a New Study — Higher Rates of Crime." *The74million.org*. //www.the74million.org/failing-schools-home-to-underachieving-students-disillusioned-teachers-and-according-to-a-new-study-higher-rates-of-crime/.

[11] Benjamin, *Race After Technology*, 81.

[12] O'Neill, *Weapons of Math Destruction,* 28.

This practice "essentially crowdsources fear, and that fear is racialized,"[13] as another form of modern-day redlining.

Often, the champions of predictive crime models that deem certain neighbors to house more illicit activity fail to question if the original arrest data is racially motivated. It has been recorded that New York state officers have been known to persuade victims to not file complaints to maintain lower stats for serious crimes, as well as falsifying other crimes to meet quotas. This intentional manipulation of data, known as "juking the stats", for crime reporting programs is their attempt create the image of a more successful jurisdiction, and colored communities are easy to target as they are already under more surveillance[14]. Additionally, when thirteen jurisdictions in New York were under investigation for these illegitimate policing practices and subsequently corrupt data, they still continued to use predictive policing tools based on such dirty data[15]. Due to subjective authority over these systems, data is becoming more a reflection of police practices instead of the information it seeks to represent. Such that 'data' may be a misnomer because data should imply "some type of consistent scientific measurement or approach," while actual crime data is often incomplete or can be distorted. There is insufficient reformation and government transparency among predictive models. Between the faulty data collection and opaqueness of algorithms, we are left to question, "what is the

[13] Daniels, J., 2019. "'Color-blindness' is a bad approach to solving bias in algorithms." *Quartz.* https://qz.com/1585645/color-blindness-is-a-bad-approach-to-solving-bias-in-algorithms/.
[14] Richardson, R., Schultz, J. and Crawford, K., 2019. "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice." *Papers.ssrn.com.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423.
[15] Ibid.

difference between over policing in minority neighborhoods and the bias of the algorithm that sent officers there?"[16]

So, although these penal system surveys conducted are not allowed to inquire race, it is simply not necessary. When bodies of color are overpoliced compared to their White counterparts and more likely to grow up in lower income communities, the rest of their data will still ensure they are racially differentiated by the questionnaires. These 'neutral' risk scores are ignorant to this reality of institutional racism[17] that creates racial disparities across all social factors. This is an example of 'punishing the poor' for the demographics they are funneled into, and they are continually punished today for their resulting impoverishment.

Due to all this potential harm, we often do not hear how the system actually works—when increased surveillance induced by predictive crime models does prevent crimes or inmates are judged more efficiently and effectively. The potential for good is a genuine reality of data models, but regularly obstructed by the abusive cases since the results of these algorithms are not light matters.

We can see the perpetuation of discriminatory practices in the case of the two young Black women in Coral Springs, Florida, who were charged for theft after playing on bikes left on the street and were subsequently predicted as more likely to commit another crime by the Northpointe questionnaire than a White man who was a seasoned offender of armed robbery. This questionnaire was composed of 137 questions, where not a single one needed to inquire about race to develop the same strata. In 2012, the Wisconsin Department of Corrections

---

[16] Raji, D., 2020. "How our data encodes systematic racism." *MIT Technology Review*. https://www.technologyreview.com/2020/12/10/1013617/racism-data-science-artificial-intelligence-ai-opinion/#:~:text=Data%20will%20always%20be%20a,models%20that%20interpret%20the%20information.

[17] Benjamin, *Race After Technology*, 81.

launched the use of this software throughout the state, from determining sentencing to parole. However, only 20 percent of the people predicted to commit violent crimes actually went on to do so. Additionally, black defendants were labeled to commit future crimes at twice the rate of white defendants[18]. Even the Wisconsin Assistant Attorney General Christine Remington fights the unwavering faith that the justice system places in the score to determine sentences–unjust disparities have become common in the criminal justice system and society alike[19]. Also, like the LSI-R, defendants risk scores are not available to the public so they cannot assess what exactly causes such severe and apparent racial bias.

It is difficult to construct a score that doesn't include items that can be correlated with race — such as poverty, joblessness and social marginalization—as accuracy goes down without these, which could negatively impact other scores.[20] It can be difficult to mitigate the harm of these social structures, as racial discrimination is interwoven in our institutions. However, we will explore some potential procedures that could alleviate the detriments of such situations.

Teacher Evaluation Model

Another model that staggers between the potential for both betterment and harm is the IMPACT evaluations developed by the District of Colombia Public Schools. This assessment of classes' math and language proficiency was created to celebrate high-performing educators and

[18] Kirchner, L., Angwin, J., Mattu, S. and Larson, J., 2016. "Machine Bias." *ProPublica.* https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
[19] "In Baltimore itself, the last ten years have resulted in federal investigations into systemic data manipulation, police corruption, falsifying police reports, and violence, including robbing residents, planting evidence, extortion, unconstitutional searches, and other corrupt practices that result in innocent people being sent to jail,"–Leading to 55 potential lawsuits in 2018 alone. Ibid.
[20] Ibid.

reform underperforming classrooms[21]. Certainly, these areas of education are important and should be recorded and analyzed. However, this algorithm is the perfect example of how data science can be manipulated at many steps of the process. Under the scope of faulty data, if a class's scores are embellished to look impressive, the following year the next teacher appears to not perform as well in comparison as the progress looks stagnant or weakening. Then they are punished for maintaining the integrity of their own class' scores. The opaqueness of this process worsens the harm because it is difficult to detect this possible downfall of the data, and we still continue apply the results regardless due to an inflated trust in mathematics. Again, the results themselves are typically helpful and are valid assessments of significant educational data, but how the results are applied can become abusive to the teachers. If the score alone is the determining factor of whether or not they are an adequate educator, despite any personal testimonies or confounding variables that affect student performance, then the score does not simply assess the quality of the teacher but determines it[22]. A model cannot analyze external factors that are vital considerations for education quality levels. This situation is an issue of the employment of models, which falls under the umbrella of poor data practices.

ETHICS OF DATA

Quality of Data

In the eye of general public, algorithms are deemed unbiased and fair to all input. However, not all data is created, collected, or processed equally. We see this possible distortion with the 'juked' statistics of the New York precincts, or exaggerated student scores in the

[21] Dcps.dc.gov. 2009. "IMPACT: The DCPS Evaluation and Feedback System for School-Based Personnel." *DCPS.* https://dcps.dc.gov/page/impact-dcps-evaluation-and-feedback-system-school-based-personnel.
[22] O'Neill, *Weapons of Math Destruction*, 4.

IMPACT teacher analysis. Data can be subject to multiple forms of manipulation at once, which makes it extremely difficult, if not impossible, for systems trained on this data to detect and separate 'good' data from 'bad' data,[23] Even a perfect model will produce inaccurate results with skewed input: the "garbage in garbage out," concept[24]. And this 'garbage out' can have severe consequences as it is often regarded a fact and applied accordingly. The harmful cycle induced by 'juked stats' worsens if we do not mitigate the harm from its origins: the methods of data collection and organization, as faulty data will continue to instigate a system that produces more inequity, because the data models are solidifying illegitimately collected and organized data as fact. And "data will always be a subjective interpretation of someone's reality,"[25] such that there is an element of control within all data processes, leaving the subjects at the mercy of the designer. However, this issue of authority falls into the arena of expert practice, which we will explore further in the ethics of practice.

Complexity and Sensitivity of Data

Since data is growing at such exorbitant rates, collection and processing has become so complex such that a technical review–to determine quality and application of data sets–requires assessments of data logic, consistency, formatting, accessibility, plausibility, quality, handling and reuse, units of measurement, quality of collection method, and if any anomalies are present[26]. However, these intricate processes are informally structured, and many scientists use data sets that have not been peer-reviewed. However, this also becomes more of an issue of faulty practice. Peer-reviewed datasets allow for increased authority and transparency in their

[23] (Richardson, Schultz, and Crawford, 2019)
[24] O'Neill, *Weapons of Math Destruction,* 150.
[25] (Raji 2020)
[26] Enago Academy. 2022. "Should Data Sets Be Peer Reviewed?" https://www.enago.com/academy/should-data-sets-be-peer-reviewed/.

data analysis, but often it is primarily articles that are peer-reviewed, instead of the original data that is shared between researchers and companies. Most data scientists champion data sharing because this data collection is expensive, in terms of funds, time, and manpower. Open data sets have immense potential for good as data scientists can work more efficiently, but if faulty data is circulated between researchers and companies, it can lead to more defective data outputs. Unfortunately, while peer-review is a common practice for articles, the process for raw data is mostly undefined[27].

Data is highly interconnected in today's market so we can see in our case studies that zip codes have far more meaning than one's location, especially with models that stratify economic groups. Due to this nature of data today, we can see deep into subjects' lives with only a couple of pieces of personal information. "When these data are linked together, they provide an electronic shadow of a person" or group, even without explicit identifiers of their person[28]. If the machine only has to identify race via basic information to potentially suppress and abuse entire communities, the casted shadows become large masses that are extremely easy to identify. And ironically, we have to sell more of ourselves away to gain privacy and protection, as with facial recognition and fingerprints to protect our information and personal objects.

Personally Identifiable Information (PII)[29] protective regulations rely on the sensitivity of the data, but when we consider race, which is both highly sensitive–given the effects of racial bias–and correlation with almost all other information, the question becomes how can it be

---

[27] Ibid.
[28] Sweeney, L., 2001. *Computational Disclosure Control: A Primer on Data Privacy Protection*. [online] Dspace.mit.edu. Available at: <https://dspace.mit.edu/handle/1721.1/8589> [Accessed 20 November 2021].
[29] "Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means." dol.gov. n.d. *Guidance on the Protection of Personal Identifiable Information*. [online] Available at: <https://www.dol.gov/general/ppii> [Accessed 5 May 2022].

handled such that it does not impact the output of algorithms. The solution to this is extremely complex and currently imprecise as this reality of interconnected data is the result of hundreds of years of racial discrimination.

There is a complex balance of privacy and properly educated systems, for sufficient refinement in data granularity is important for the decision-making process[30]. So, simultaneously, lower income groups rely on refined data for algorithms to meet their specific needs, say with government provisions, but are also the subjects of oversurveillance and privacy breaches in highly refined data collection and analysis. Because of this interconnected nature of race with wealth, education, location, we encounter the issue of data granularity that cannot ever be reduced enough to fully prevent racial differentiation.

Regardless of this particular issue of interdependent racial demographics, the perfect equilibrium for all data between benefit and risk is unachievable, as it is highly subjective. For example, there is significant benefit in using large-scale medical records to infer conclusions on effective disease treatment. However, extremely sensitive data is necessary to make these conclusions, such that privacy is at risk. There is not a flawless level of data collection to be met, yet we recognize that not all data is created equal, and subsequently, faulty data has been permitted without sufficient peer-review to detect it.

ETHICS OF ALGORITHMS

Imperfect Models

When we employ non sentient machines to sort through millions of datum, it appears that human bias is completely removed and the results should be objective. However, the machines

[30] Bello, W., 2022. *Pitfalls in privacy data wrangling*.

need not make conscious decisions to discriminate against certain groups when the data already reflects their disadvantages and institutional racism. Additionally, as the infamous statistician G. E. P. Box remarked, "All models are wrong; some are useful."[31] Models are not trustworthy and authoritative simply because they employ complex mathematics. They are often simplistic representations of real-life scenarios, their potential application—and potential harm—are extremely complicated as data science is ubiquitous and affects most areas of our lives, which can evoke the suffering of communities who are overexposed and hypersensitive to their results[32]. If we examine data models with the mindset of guaranteed imperfection and possibly haphazard design, given the sheer number of models throughout our daily lives, we can observe that it is not bizarre to find a model that is not only wrong and useless, but also harmful.

American mathematician and data scientist Cathy O'Neill coined the term 'weapon of math destruction' to describe the downfalls of faulty data science, which are typically the algorithms built with inherent bias or obscured in a 'black box' of data science, such that the internal processes are concealed. These attributes can lead to the abuse of the communities who are already vulnerable, adding insult to the injury of generational socioeconomic disparity– specifically toward people of color.

The intentions of WMD (weapons of math destruction) are typically not malicious and actually seek to improve the inner workings of society; we typically believe if we replace the potentially biased mortgage brokers and hiring managers of the past with neutral models, we will

---

[31] Barroso, G., 2019. *"All models are wrong, but some are useful". George E. P. Box – AdMoRe ITN*. [online] Lacan.upc.edu. Available at: <https://www.lacan.upc.edu/admoreWeb/2018/05/all-models-are-wrong-but-some-are-useful-george-e-p-box/> [Accessed 30 March 2022].

[32] Hand, D., 2018. *Aspects of Data Ethics in a Changing World: Where Are We Now? | Big Data*. [online] Liebertpub. Available at: <https://www.liebertpub.com/doi/full/10.1089/big.2018.0083> [Accessed 15 November 2021].

open the playing field. But we cannot always be concerned with equality of models, but equity--allocating the knowledge and sensitivity needed to fairly encompass all groups. We do not all have the same starting line, as seen by the persistent implications of systemic discrimination of certain groups. And we question, how can a model measure a situation like a human can, weighing it individually and situationally, while also avoiding human bias? It is currently implausible to facilitate both of these attributes given the current limits of artificial intelligence.

This indulgence in an inflated epistemological confidence in a mathematical model will often transfer that confidence onto the resulting conclusions of the model[33]. However, models are essentially significantly simplified versions of human analysis, which is already highly biased due to racially skewed institutions. Then, when models lack far more contextualization and situational awareness than their human counterparts, we have to realize that these algorithms are not always dependable. We consider once emotion is removed in an algorithm, we will reach peak reason and logic, and thus equality for all subjects. However, machines cannot always be trusted to consume and process all information equally when designed by humans with their own inherent bias.

False Neutrality of Models

Racism can be amplified through algorithms because they not only take in racially charged data into a bias model, but then apply it as object fact, all while claiming its neutrality—embodying the "digital denial"[34]. Even "raw data reflect deeply ingrained cultural prejudices and structural hierarchies,"[35] such that results can perpetuate harmful marginalization. Thus,

---

[33] Howell, R. and Bradley, J., 2011. *Mathematics Through the Eyes of Faith*. New York: HarperOne, p.216.
[34] Benjamin, *Race After Technology*, 12.
[35] Ibid, 58.

algorithmic outputs cannot consistently sustain this belief of technological innovation directly inducing societal advancement.

It is a growing belief that a machine will be skewed if the model itself is 'undereducated' by being colorblind. Historically, data scientists considered a colorblind model to be the unprejudiced and unbiased approach, but when racial divisions are inherently tied to a majority other demographic stratification—like education level or zip codes–a model will still find methods to separate subjects into racially "damning buckets"[36].

The proponents of race comprehension strongly believe that shielding AI from the 'bad' side of society leads to an imbalance—a deprivation of the knowledge of racial literacy. Such that if we expect them to make the decisions of man, they need as many similar complexities of the mind that can be reproduced. As social colorblindness can lead people more ignorant to the needs and cultures of unique communities and individuals, likewise, colorblind algorithms could lead to more ignorant outputs. We trust these algorithms to process data and produce results that will affect every area of our life, but without expecting these to be taught the vital components of a racially charged society.

To combat racial associations in algorithms, the technological industry has been known to utilize implicit association tests (IAT)[37] in bias trainings, which measure the strength of associations, evaluations, and stereotypes in their model designs[38]. Furthermore, a project called TakeTwo has emerged to challenge racial language in writing and design, by detecting words and phrases that could be perceived as racially biased to help content creators mitigate harm.

---

[36] (Raji 2020)

[37] IAT originated to assess and teach inherent gender bias in STEM because previous diversity trainings did not improve tendencies towards stereotypes of women in these disciplines. Jackson, Sarah, Amy Hillard, and Tamera Schneider. 2013. "Using Implicit Bias Training to Improve Attitudes Toward Women In STEM". *Social Psychology of Education* 17 (3): 419-438. doi:10.1007/s11218-014-9259-5.

[38] (Daniels 2019)

This is a small step towards recognizing implicit bias, but this model could be translated to multiple datasets that utilize phrases in addition to variables[39]. Unfortunately, in general, there is currently insufficient experimentation with racially literate algorithms despite the immense investigation into the false notion of truly colorblind models. There is hope that these experiments will have to emerge given the palpable nature of racial skew within data.

It is not the notion of a colorblind model that is the issue; obviously we want to remove the confounding variable of race to analyze communities equally. It is the complexity and interconnectedness of data that makes true colorblindness exceedingly difficult and has forced many data scientists to analyze the possibility of informing their models of the skewed effects of race in modern institutions. Models are themselves neutral to the information and formula it is given, but when the infrastructure is corrupted, the whole system is can be compromised so that equality does not ensure equity. Algorithms can only process data in the manner they were taught—or not taught. It cannot take into account the complex elements of disparities among races when it was never programmed to do so. Remaining neutral may be unrealistic and allowing such gaps in the knowledge of data science teams and their machines leaves them to develop their own biases based on socioeconomic situations among similar communities that is reflected in most data.

For example, surnames are very easily ethnically distinctive, while many first names are also often racially coded. White culture is typically seen as the norm or baseline, even though it is its own unique culture and does not make up the global majority[40]. Nevertheless, those outside

---

[39] Srivastava, P., 2022. "We must check for racial bias in our machine learning models." *IBM*. https://www.ibm.com/blogs/journey-to-ai/2022/02/we-must-check-for-racial-bias-in-our-machine-learning-models/.

[40] (Raji 2020)

of this imaginary norm are still processed as different. There is a classism tied to ethnically unique names, such that there is a safety in the 'blandness' of names that 'sound' White. There is no neutral culture, but Whiteness is typically set as the standard of acceptability. Names that are typically used in ethnic communities are more often pulled for airport screening, police frisking, and rejected from job applications due to both analog and automated systems[41]. Even our search engines correlate certain ethnic-sounding names with discriminatory stereotypes, such as a search for gorilla generating images of Black people[42]. When Princeton data scientists allowed an algorithm to categorize between White and Black names, the White names were associated with pleasant words while Black names had the opposite correlation[43].

Hiring Algorithms

Each subject has their independent information, such that the actions of an individual should not affect the perception and analysis of others with the same or similar names, yet skewed models can work otherwise. For example, despite 'colorblind' hiring algorithms, "job seekers with White-sounding first names received 50 percent more callbacks from employers than job seekers with Black-sounding names persisted across occupations, industry, employer size,"[44] which facilitates the cyclical nature of poverty among people of color. For a time at Amazon, their hiring algorithms utilized in-house data of employees—majorly made up of white

[41] Benjamin, *Race After Technology,* 4.
[42] (Daniels 2019)
[43] Benjamin, *Race After Technology*, 5.
[44] Ibid.

men—to judge the next application pool. This led to less hiring opportunities for women and people of color, even if they had the same educational backgrounds as the current employees[45].

Before the actual application process even begins, the advertisements of job opportunities are typically generated from personal internet data, which further divides in the workforce[46] as individuals who are still 'outside' of a given discipline will have less internet history regarding that field and also may fall into racialized and gendered 'buckets' because of said history. Then the application pool is processed by hiring algorithms, in which job seekers are completely shielded from the manner in which they are assessed and thus clueless to what removes them from the group of potential candidates.

Subsequently, applicants seek to learn to play the system, such as using white text to conceal keywords of application requirements or names of highly ranked universities so hiring algorithms would still register their resume as favorable[47]. Ethically speaking, even if these applicants are making the system more skewed, this shift towards the other direction could be considered their attempt at equity as they are simply trying to level the playing field of an already twisted system. Companies have been known to record 'personality' questionnaires, which act as proxies[48] for unlawful mental health questionnaires[49], in which applicants also may lie to look agreeable. When these unethical loopholes can be justified for hiring, it makes sense

---

[45] Poduska, J., 2021. "You've Been Warned–Bad Data Models are Capable of Destroying Companies." *Datanami*. https://www.datanami.com/2021/11/29/youve-been-warned-bad-data-models-are-capable-of-destroying-companies/.
[46] Bogen, M., 2019. "All the Ways Hiring Algorithms Can Introduce Bias." *Harvard Business Review*. https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias.
[47] Buranyi, S., 2018. "How to persuade a robot that you should get the job." *The Guardian*. https://www.theguardian.com/technology/2018/mar/04/robots-screen-candidates-for-jobs-artificial-intelligence.
[48] The practice of proxies in data science are meant to stand in for immeasurable variables that are highly correlated.
[49] O'Neill, *Weapons of Math Destruction,* 106.

that employees are warping their answers, because these answers should not determine their career capabilities in the first place.

Credit scores have also been used to determine job capability. And although they cannot mandate the submission of this sensitive information, this simply is the mirage of consent, for those who do not comply may not be considered at all. For those in lower income communities, not only do they already have less favorable job history due to lesser educational opportunities, but likely have lower credit rates as well.

Of course, algorithms sort through applications more efficiently, but the notion of effectiveness is subjective. Should a company be considered more effective because they hire in less time for less cost, or because they provide equal opportunities for all communities in their field by intentionally considering their applicants? They are not only harming vulnerable communities but harming their own company's ability to perform by restricting the opportunity to hire diverse employees.

Invented Notion of Algorithmic 'Glitches'

These automated systems are attractive not only for fiscal efficiency, but also because it seems to remove human culpability, hence the scapegoat of the 'glitch' of machines. This captures Benjamin's concept of the "allure of objectivity without public accountability."[50] Models can lead us to move from dynamic action to passive action—to save time, money, and responsibility—as less humans are faced with unique decisions that come with interacting with and processing individual information, say with loan approvals or job applications. Instead, we allow predictive technology to decide the fate of subjects. Like the IMPACT scores, which had

[50] Benjamin, *Race After Technology*, 52.

been previously used to establish the quality level of the teacher instead of simply acting as a supplemental assessment alongside personal testimonies. These specific methods of model implementation fall under the umbrella of the ethics of practice, because an evaluation score is helpful until it is used to embody the full picture of a teacher's career.

Often, glitches are simply marked as outliers, such as two first offending young black women deemed riskier than the white male seasoned offender. However, glitches are not simply infrequent errors, but consistent corruption of designs that have come to be accepted as social norms. These women were clearly not assessed independently based on their individual actions, but by the demographics that separated them into a particular 'bucket', thus bearing the weight of the action and consequences of everyone else in their 'tribe'. Real data should be concerned with the specific participants at hand when the outcome can have such harsh ramifications, especially when said individualized data is so readily available today. Similar data points are not necessarily harmful when analyzed in general, but become dangerous given the intended use, like determining criminal punishment.

Credit Assessments

The arenas of data algorithms and practice often bleed together, but we can still identify models that inherently have flaws. Take Zestfinance's alternative risk assessments using language skills to determine borrowing limits.[51] This model can be extremely helpful to some communities, but if a potential customer is lacking in these—which often points to poor education, thus, to class and race—it can punish borrowers for their institutional disadvantage.

[51] "Their goal is to bypass credit scores for those who do not have favorable credit. While gauging credibility is unusual observations like punctuation, capitalization, reading speed, and more." O'Neill, *Weapons of Math Destruction*, 158.

Of course, this model is a voluntary risk assessment, but many major companies, like AMEX and Capital One, also have been known to employ skewed models or practices. Therefore, the viable options for those lower on the socioeconomic ladder significantly decreases when the trusted companies employ faulty models.

Capital One created a scoring system to mimic credit scores to gauge prospective borrowers' wealth, which cross references their locations to determine their economic status. Since they do not have the exact data they are trying to account for when generating advertisements, they use proxies of shopping history to fill these gaps, creating 'e-scores'. Those who land a low score on this Capital One's proxy scale receive advertisements for credit cards with lower credit limits and higher interest rates, enforcing another malicious feedback loop for those of within the lower income demographic.[52]

Similarly, AMEX has been known to cut spending limits on customers who shop at particular 'lower-income' stores, such as Walmart, thus lowering their credit and driving up borrowing costs[53]. This is another instance where they essentially create the situations the lenders are trying to avoid: by lowering impoverished borrowers' limits, they are more likely to violate their limits and be fined, and then have less money to pay off their fees, leading to further decreased limits and increased rates. It is important to acknowledge, yes, lower income borrowers are riskier on average, but creating situations that worsen their ability to borrow effectively is the fault of the lenders' practice, not the borrowers'. Human agency should be maintained in some form, so that there is an opportunity to appeal to human rational and situational awareness. Instead of solely relying on the black and white decision of machinery, especially when bizarre circumstances cannot be accounted by algorithms. Thus, the results of

[52] O'Neill, *Weapons of Math Destruction*, 143.
[53] Ibid, 156.

most models should be assessed before taken solely as definitive fact or accompanied by supplemental human decision.

Intention of Applications

Seemingly harmless data becomes dangerous under specific circumstances. Like online shopping history; one algorithm tailors advertisements to a user, while another algorithm uses the same information to determine their credit card spending limits, simply based on the sites and companies they purchase from. The arenas of algorithms and practice begin to bleed together, as these skewed models are the results of poor practice.

Niche marketing can actually evoke more stereotypical containment as more abstract characteristics like race, ethnicity, and gender are quantified for profit[54] by abusing their unique vulnerabilities. Also, the categorizing of buyers into similar socioeconomic and educational buckets may coerce them into the same actions simply from the advertisements they receive.

This example relates to Princeton professor Ruha Benjamin's question of "what do 'free will' and 'autonomy' mean in a world in which algorithms are tracking, predicting, and persuading us at every turn?"[55] And for vulnerable communities, are they provided abundant choices like the privileged, or funneled into specific options due to the technology that reinforces such socioeconomic standing? As identity has always been shaped by outside forces around us, this is just a new form of molding, one with less opportunity for those marginalized.

Because of the general assumption that all technological evolution is undeniably correlated with societal betterment, technology has been deemed an unparalleled authority of knowledge, especially given its complex structures and widespread integration. This belief

---

[54] Benjamin, *Race After Technology*, 21.
[55] Ibid, 31.

mystifies technology, as its systems are typically not common knowledge, especially when specific systems are intentionally obscured from the public. The labyrinths of mathematics make up the 'black box' of algorithms, mystifying the hidden functions of data technology[56] and gains our trust because we are unaware of their true functions. The hiring managers and judges may also be under the illusion of pure advancement, as their difficult decisions are made for them by machines, while job applicants and inmates experience succumb further to vicious feedback loops because they do not have the extensive capability to reprimand such obscured practices[57].

Additionally, we cannot judge artificial intelligence by its intention or seemingly unbiased nature, but the impact of their executions. Companies should be required to investigate and peer-review their input and output data, as well as their model designs before relying on them to sufficiently calculate reliable results. We cannot make it the job of artificial intelligence to modernize and evolve alone, because it is unreliable to constantly use code that does not take into account the shifting fluidity of our world.

Dynamic data always demands dynamic models— "Conditions change, so must the model."[58] Culturally involved data is constantly changing, so models cannot remain static and uncultured to our reality, whether they are colorblind or not. AI can develop knowledge, especially via deep learning design[59], but it cannot be depended upon for institutional change, as that relies on the practice of Man. We want to trust algorithms but first we must teach and verify

---

[56] O'Neill, *Weapons of Math Destruction,* 134
[57] Benjamin, *Race After Technology*, 141.
[58] O'Neill, *Weapons of Math Destruction,* 18
[59] "Deep learning is a type of machine learning and artificial intelligence (AI) that imitates the way humans gain certain types of knowledge." Burns, Ed, and Kate Brush. 2021. "What Is Deep Learning and How Does It Work?". *Searchenterpriseai*. https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network.

them. The intentionality of sensitive models may not only be more careful to avoid institutional racism, but also help developers themselves to expand their awareness of issues potentially obscured by their privilege.

There is never a possibly benefit from data if we do not explore it, and if we do not take creative risks in our models, we cannot explore it. Additionally, we know that subconscious bias and vicious feedback loops are rarely from malicious design. Nevertheless, the sentiments of good intentions are lost when the outcomes of models are detrimental. Isolated, good models can have potential for improvement, but when manipulated by poor practice, they turn into arenas of damage.

ETHICS OF PRACTICE

Every stage of data development affects the process as whole and therefore requires intentional execution of practice. The way our data is collected affects how the data is tidied and then subsequently analyzed. The analysis level is intricately composed, as it relies heavily on the data scientists themselves to determine the regulations and many procedures of their practice.

Lack of Diversity

Although the bias discussed above can be difficult to eradicate in artificial intelligence, reinforcing racial literacy among data scientists is always vital, so that they are equipped to ask how preexisting institutional racism will interact with their design. As well as awareness to ask if their teams have diverse enough perspectives to recognize potentially racially skewed elements. It is not that the ethnically diverse data scientists do not exist, but that they have a harder time getting hired on average. According to the United States Census Bureau, Hispanics and Black

communities do make up small percentages of STEM graduates, that of eight percent Hispanic and six percent Black, but the diversity reports of major tech reflect even lower numbers on average for their employment: only three percent Hispanic and one percent Black.[60] For data science courses through General Assembly[61], the amount of Black and Latino students has the lowest enrollment percentage compared to other technical disciplines like web design and digital marketing—the average Black and Latinx enrollment across all technical concentrations was 17.5 percent while only 11.8 percent was recorded for the concentration of data science specifically. Thus, "Data Science seems to draw from a smaller, more specialized pool, which could, in part, perpetuate diversity issues."[62] Forbes also claims that Black and Latinx high students are afforded less opportunities for higher level math and science classes, which likely is due to lack of public-school funding in lower income areas, as previously discussed.

A decent amount of the faultiness of data practices surrounding race is likely due to this stratification, not malintent. Those dominating the field of data science may be unaware of not only the intensity of their impact in general, but how it specifically affects the marginalized communities as they lack insider perspectives. Although bias can be difficult to eradicate in artificial intelligence, reinforcing racial literacy among data scientists is always vital, so that they know how to ask how preexisting institutional racism will interact with their design and if their teams have diverse enough perspectives to recognize potentially racially skewed elements.

Beyond this trend, we will explore the other areas of practices in which data science can begin to deteriorate.

---

[60] Landivar, Liana. 2013. "Disparities In STEM Employment By Sex, Race, And Hispanic Origin". ACS. *United States Census Bureau*. https://www.census.gov/library/publications/2013/acs/acs-24.html.
[61] "an education company that trains students in data science and other technical fields." "The Data Science Diversity Gap". 2017. *Forbes*. https://www.forbes.com/sites/priceonomics/2017/09/28/the-data-science-diversity-gap/?sh=3db53fde5f58.
[62] Ibid.

Handling of Datasets

When the importance of intention is depreciated, as can happen with the use of open data sets and purchased data, the information loses the supportive structure of contextualization and abuse can ensue. When possible, if data scientists use raw original data in their publication, it facilitates transparency when examining their intentions. However, this is not always possible, because otherwise we would waste value data that already exists and that could be utilized to help people now. Nevertheless, datasets acquired for the purposes of research should not be treated the same as data aimed for institutional operations, as the latter attempts to solidify data as fact.[63] We cannot apply correlations we are still trying to form and prove through research.

Data sharing can allow one stigma of a person to be passed around into all areas of his existence. If data scientists do not intentionally investigate the borrowed data, they may have no way of knowing if the data is corrupted or biased. For example, the peer-to-peer system[64] allows lenders to make and share their own regulated models—that they do not have to disclose to borrowers—like the creation of fake credit scores based on salary, occupation, debt, shopping history, etc.

An example of dangerous data storage is California's gang database, created as an attempt to facilitate social control. This database was created through a combination of zip codes and racialized names; thus, it was majorly composed of Blacks and Latinos, even including babies younger than a year old. Once you are categorized, it is hard to alter your standing—databases are often not destroyed and are open for further use, as these California gang databases

---

[63] (Hand 2018)

[64] "Started out in the last decade with the vision of borrowers and lenders finding each other on matchmaking platforms…the Lending Club…generated e-scores, which they claimed were more accurate than credit scores." O'Neill, *Weapons of Math Destruction,* 159.

were not scheduled for liquidation for more than 100 years[65]. "Big data seems to be fostering the false notion that we have an obligation to retain any data that we come across because of its potential usefulness,"[66] but data should expire as peoples' situations are always changing, and corresponding new data is being collected constantly so that older data may not be as consistently useful as we hope.

Data Regulation

Data is always growing and evolving, and there is still dramatic change to come in regard to the possible application of data science, such that it is complicated to properly regulate data science because creating and maintaining precise regulations is far-fetched. "Principles must be mapped to low-level guidance, and this is likely to be application specific,"[67] as general regulations cannot capture all the complexities of the field of data today. While technological innovation is constant, society lags to adjust accordingly, especially regarding legal policies.

Data professionals commonly want to establish their own moral regulations to maintain autonomy from external control, stressing the protecting of free speech in data. Companies simultaneously reassure employees that they work for a trustworthy organization and the public that they are protecting and benefiting them, but these statements are unsurprisingly vague, and practices vary from company to company.[68] Companies implement hands-off policies to protect their intentions and application, which expands the opaqueness of the 'black box'. The allowance of proxies also facilitates accelerated analysis and increased financial turnover, such that the

---

[65] Benjamin, *Race After Technology*, 7.
[66] Harris, Jim. 2013. "Data Has an Expiration Date". *The Data Roundtable*. https://blogs.sas.com/content/datamanagement/2013/01/16/data-has-an-expiration-date/.
[67] (Hand 2018)
[68] Ibid.

"push towards privatization is at the expense of other human needs,"[69] because their self-regulated, obscured models are means to a profit.

In general, data privacy is not highly protected; as we see in the US, there is no single law encompassing the protection of personal data. There are specific sectors of data regulations, as with HIPPA or FERPA, but a majority of companies are left to their own devices to self-regulate their privacy practices due to lack of rigid data regulations[70]. Private companies are allowed to decide which data are valuable and deterministic, what goals to optimize for, and who is affected by their analysis. However, like any other social system, it requires regulation. The lack of federal and state data policy is abysmal considering the effect data science has in our lives daily, in our consumption, borrowing, residing, and much more. The lack of data security and lag of the government is sharply felt, and according to *Race After Technology*, 55% of their public sample want more regulations[71].

The United States' Personal Data Privacy and Security Act of 2009 never passed, which would have not only enhanced the punishment for identity theft, but also protected against misuse of personally identifiable information[72]. However, we do have to acknowledge some of the advancements as of late in data regulation. The EU has seen a majority of the mass advancements, such as with the General Data Privacy Regulation (GDPR) which enforces transparent data processing, minimizing data collection, maintaining accurate data,

---

[69] Benjamin, *Race After Technology*, 29.
[70] Klosowski, Thorin. 2021. "The State of Consumer Data Privacy Laws in The US (And Why It Matters)". *Wirecutter: Reviews for The Real World*. https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.
[71] Benjamin, *Race After Technology*, 38.
[72] *Personal Data Privacy and Security Act Of 2009*. 2009. Vol. 1490. 111th Congress.

accountability of data controllers, time limits of data storage, and more.[73] In the US, some state governments have enacted their own data policies, as with California's Consumer Protection Act which allows residents to ask companies to delete data or prevent the sale of their data[74]. Nonetheless, state policy is a slow spark across the whole of the US, thus why companies are largely left to self-regulation. With profit on the line and incredibly weak barriers of their operations, self-regulation has simply led to the mirage of protected data and opaque data analysis and implementation. Companies are unlikely to forgo profit over individual data protection, especially if they can maintain the 'black box' when the United States does not have its own GDPR.

In the way of data policy and encryption today, it has been discovered that not long-ago law enforcement leaders like FBI Directory James Comey requested that technology companies create "backdoors" in their products to allow for government investigation of users. And since we lack federal regulations—by the government's own design—this is perfectly legal within the structure of self-regulated companies.[75]

Authority of Data Practice

So, as it is the responsibility of data scientists to source and select data, as well as design and develop models[76], they hold an extensive power as the authoritative experts of data. They

---

[73] Violators of these principles will face harsh fines and potential lawsuits of violated subjects. Wolford, Ben. 2022. "What Is GDPR, The EU'S New Data Protection Law?". *GDPR.Eu*. Accessed January 10. https://gdpr.eu/what-is-gdpr/.

[74] Stuart Thompson, and Warzel, Charlie. 2019. "Opinion | Twelve Million Phones, One Dataset, Zero Privacy (Published 2019)". *Nytimes.Com*. https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html.

[75] Bedoya, Alvaro. 2016. "What the FBI'S Surveillance of Martin Luther King Tells Us About The Modern Spy Era". *Slate Magazine*. https://slate.com/technology/2016/01/what-the-fbis-surveillance-of-martin-luther-king-says-about-modern-spying.html.

[76] (Raji 2020)

can use their models to create more biased results if their practices are distorted and/or have obtained inadequate data. The vicious feedback loop continues, especially if data is transferred between companies as it often is. If an inmate is denied parole because of racially charged questionnaires, it will then certainly go on his record and his extended sentence will affect his next parole application, his future job applications, future renters' applications—thus almost every significant arena in his life.

As we see with the LSI-R predictive model, many criminal risk assessments are sourced from private companies, who are almost entirely controlled by self-regulation alone. "Validity had only been examined in one or two studies" and that "frequently, those investigations were completed by the same people who developed the instrument."[77] Researchers Sarah Desmarais and Jay Singh could not find any substantial set of studies conducted in the United States that examined whether risk scores were racially biased— "The data do not exist."[78] So, it has become acceptable to bypass peer review and case studies to confirm accuracy before application.

New York State started using the predictive crime tools in 2001 to assess people on probation, and by 2010 had already implemented it in almost all other probation departments. However, "the state didn't publish a comprehensive statistical evaluation of the tool until 2012,"[79] which was perfectly legal, even though frowned upon. Peer review and evaluation of models would significantly improve underperformance of analysis, which could also help jurisdictions process subjects more effectively. Companies and governments should never implement a data model without ensuring its ethicality, as the New York justice system did. The crimes of data abuse will end up on their hands, as we no longer can excuse the glitch of systems

[77] (Kirchner, Angwin, Mattu, and Larson 2016)
[78] Ibid.
[79] Ibid.

for human error and lack of moral thoroughness. The 'glitch' no longer exists as a scapegoat in an algorithm that is created and controlled by Man. Data is the epitome of interdisciplinary arenas; thus, the technical experts need to be held ethically accountable for their models that impact lives tremendously.

Black Box of Practices and Surveillance

When major companies are not legally mandated to disclose data practices and thus not compelled to, this also adds to the 'black box' of data. The data scientists behind the complicated curtains of technology affect every arena of our lives. Transparency of data science is uncommon because it is not federally enforced and could expose faulty models that prioritize efficiency and profit over equity. Facebook championed the idea of "move fast and break things", did they care what–or who–may be broken in the process?[80]

While we consent to data collection daily, we severely lack informed consent. They demand transparency of our data, but there is no two-way agreement; it has become socially acceptable to both conceal their data practices while overexposing vulnerable groups. Existing privacy statements utilize elevated jargon, typically beyond layman's terms. We must also consider the sheer quantity of consent prompts users encounter everyday—the 'common man' simply does not have the time to expend on educating himself with safe data practices while trying to simply navigate his phone. For example, when almost every cellular application request location tracking, most users likely allow it without a second thought. They may not even know why they should deny this request, because the public has been trained to believe technology should be improving and protecting our lives. When technology has become so mass spread that

[80] Benjamin, *Race After Technology,* 12.

it so natural and inherent in our society, that it becomes increasingly difficult to question these processes as we do not detect them as abnormal.

Consider the location software enveloped within our daily smartphone applications–only a finite amount of people will travel to the exact places your personal phone does, especially near your residence and place of work. The software development kits (SDKs) used for tracking are integrated into almost every application, from a local news app to using your search engine to find coupons.[81] This data is used and circulated by dozens of companies constantly. US citizens do not want the government to track and scrutinize their every move like the extreme discriminatory surveillance of the Uyghur community in China[82], yet they allow it to be privatized among companies that cannot be held as accountable as our government[83]. While the Uyghurs situation is extreme, this is not a farfetched reality of a dystopian narrative—this technology is readily available so potential harm is ballooning alongside technological advancements and poorly regulated practices.[84]

The simple promises of companies only sharing data with 'vetted partners' should not entice our corporate faith, as this data is never truly anonymous or truly consensual. "The greatest trick technology companies ever played was persuading society to surveil itself."[85] So,

---

[81] (Thompson and Warzel 2019)

[82] The Uyghurs are an ethnic Muslim community that have been routinely forced into concentration camps as an attempt to force their assimilation of culture. After their release, they enter the 'virtual cage' of 24 hour live camera feeds with highly developed facial recognition. All of their personal information is gathered and tracked, such as educational history, government records, identification numbers, religious practices, and more. If any of their behavior is deemed suspicious, and often is, they are subjected to 'virtual fences' that further hinder their freedoms but restricting the areas they are allowed to travel to. Buckley, Chris, and Paul Mozur. 2019. "How China Uses High-Tech Surveillance to Subdue Minorities". *Nytimes.Com*. https://www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang.html.

[83] (Thompson and Warzel 2019)

[84] An independent researcher of Uyghur surveillance, Adrian Zenz, stated "It is the combination of manpower and technology that makes the 21st-century police state so powerful." Ibid.

[85] (Thompson and Warzel 2019)

now to participate fully within society we are forced to sell a little bit of ourselves, even unknowingly.

For those who have the "choice" to opt out of oversurveillance, often retention of privacy is equated with suspicious activity for lower-income communities, instead of the exercise of rights. So, privacy has reached the inevitable deterioration—to be an active member of society, data consent is really a formality because "attempts to opt out of tech-mediated life can itself become criminalized, as it threatens the digitals orders of things. Analog is antisocial."[86] We see how in India for those who refuse to comply to the Aadhaar[87] card system, are practically treated as criminals as they are denied basic services if they want to maintain their privacy of intimate personal information.

Circling back to US surveillance, we have to examine immigrant communities who are typically among the most impoverished and the most heavily surveilled due to their citizenship status. Due to this level of investigation and potential punishment like deportation and unlawful arrests, they are less likely to want to supply their information for necessary programs, such as food stamps or health care.[88] Especially when the Immigration and Customs Enforcement agency (ICE) has a history of unlawful surveillance such as obtaining the lists of Motel 6 guests and targeting those with "Latino-sounding" names. Also, during President Donald Trump's

---

[86] Benjamin, *Race After Technology*, 19.

[87] Aaadhaar cards are a government created biometric identification system, which was originally designed with the ability to opt out. However, now citizens are denied access to marriage licenses, healthcare, scholarships, government subsidies, pensions, insurance, banking, and more if they refuse to conform. Information that is collected include fingerprints, facial images, and iris scans, which is beyond the scope of details needed to provide government services. "India: Identification Project Threatens Rights". 2018. *Human Rights Watch*. https://www.hrw.org/news/2018/01/13/india-identification-project-threatens-rights#.

[88] Madden, Mary. 2019. "Opinion | The Devastating Consequences of Being Poor In The Digital Age (Published 2019)". *Nytimes.Com*. https://www.nytimes.com/2019/04/25/opinion/privacy-poverty.html.

"Zero Tolerance" campaign against illegal immigration, facial recognition and social media surveillance were used to expose offenders.[89]

Weakened Data Privacy of Vulnerable Communities

A lot of the issues come down to the exploitation of desperation. Models capitalize on desperation—luring the poor to give up vulnerable data for convenience or discounts "while the rest of us barely notice the abuse."[90] Their privacy breach typically comes in the forms of the requirement to submit immense amounts of information for government benefits or the over-policing of their neighborhoods due to underlying racist practices. This goes beyond the issues of malicious attacks of their data, because in the supposedly trusted arenas of government and privatized companies they are overexposed, due to pervasive networks of cameras, over-policing in lower income zip codes, and the location tracking.[91] Since vulnerable, lower income communities are expected to forgo any and all data to receive any government benefits, they are never provided a legitimate option to opt-out of data collection. Often, they may be completely unaware that their data is being used against them, as data is sold between covertly companies.

They also tend to fall behind the technological curve more if they experience one overwhelming privacy breach, such as a virus or hacking attempt, because they cannot afford new hardware or security. Typically, they must rely on smartphones alone for digital access, and we have seen how easily location and sensitive data is tracked via cellar devices.

So, not only are they already subject to imbalanced privacy breaches as they are watched more heavily, additionally, they do not have access to the same technological means for data

---

[89] Ibid.
[90] O'Neill, *Weapons of Math Destruction,* 168.
[91] (Thompson and Warzel 2019)

protection. This combination of general surveillance of technology users with their increased exposure is an amalgamation for immense vulnerability.

VALUE OF BODIES

Modern notions of racism often evoke the ideas of malicious intent or extreme cases of discrimination such as in Dr. Hoffman's assessment of insurability of Black citizens. Yet racism still lurks in the consciousness of man, within implicit bias due to the institutional racism interwoven throughout our society. Allowing even the rhetoric of racial systems, such as Trump's Muslim database,[92] shapes the way people think, to the point of implicit bias.

For experts in the field of ethnic and racial studies, like Black Studies professor Christina Sharpe, they see the current 'climate' or 'weather' of our society as anti-Black. While individual hate may be more sporadic and be socially rejected, institutional indifference determines the fluidity of social structures and is harder to alter significantly.

Modernity encompasses the dichotomy of valuable and devalued vulnerable bodies—the information of their communities, which are typically made up of people of color, is highly valued and overexposed, while their bodies and their persons are devalued and not truly perceived, which embodies the idea of 'those who are watched but not seen.'[93] When people are quantified it is easier to analyze them simply as data points. The 'intellectual property' of this information is highly valuable and profitable when sold within the professional sector, while

[92] "Trump administration-in-waiting is considering reinstating a database of immigrants from Muslim-majority countries — something the federal government did from 2002 to 2011." Lind, Dara. 2016. "Donald Trump's Proposed "Muslim Registry," Explained". *Vox*. https://www.vox.com/policy-and-politics/2016/11/16/13649764/trump-muslim-register-database.
[93] Chuen, Lorraine. 2018. "Watched and Not Seen: Tech, Power, And Dehumanization". *Guts Magazine*. http://gutsmagazine.ca/watched-and-not-seen/.

their tailored algorithms are also worth billions of dollars[94]. This is all heavily defended and hidden from the public to ensure the preservation of their practices. The government and companies alike are allowed to scrutinize the personal information of vulnerable communities for their own purposes, without having to fully acknowledge their humanity.

So, for us not overwhelmingly affected by poor data science practice, their extreme abusiveness evades our priorities, and is consciously ignored by companies who profit off of it. "Hierarchy that allows some modicum of informed refusal are the very top. For the rest of us…not an option."[95] The retention of privacy is seen as a privilege, not a fundamental right of all communities.

"Race itself is a kind of technology—one designed to separate, stratify, and sanctify the many forms of injustice experienced by members of racialized groups."[96] It is a social tool—a model that takes in data that is already biased and outputs further manipulated conclusions that circulate and simultaneously worsen our perspectives of vulnerable groups and their socioeconomic situations. Racial technology did not arise with data technology; long ago we allowed this racial standard to become a cornerstone of human engagement and decision making: a "tool to denigrate, endanger, and exploit non-White."[97] Now we can quantify and analyze the inferiority of colored bodies. As design is optimized, racial bias slips into the seams of algorithms[98]. Jim Crow laws were the set of legal, building, and social codes meant to maintain racial segregation, so the New Jim Code, is considered the "new racial caste system".[99] This

---

[94] Benjamin, *Race After Technology*, 29.
[95] Ibid, 16.
[96] Ibid, 36.
[97] Ibid, 17.
[98] Ibid, 10.
[99] "The employment of new technologies that reflect and reproduce existing inequities but that are promoted and perceived as more objective or progressive than the discriminatory systems of a previous era." Benjamin, *Race After Technology,* 5.

stratifies people into stigmatized buckets because of an unfaltering confidence of colorblindness and the reality of data manipulation. Individual people are forced into stereotypical categories of race that may falsify their true ethnicities since the New Jim Code seeks to achieve specialized targeting. "Criminal in this era, is code for Black, but also poor, immigrant, second-class, disposable, unwanted, detritus,"[100] and so, the vulnerable are further marginalized by poor data practices.

Their members become expendable, as long as their data is retained for profit. It is difficult to imagine a self-regulated data market in which the public does not become collateral damage. "The people running the WMDs don't dwell on those errors. Their feedback is money, which is also their incentive. Their systems are engineered to gobble up more data and fine-tune their analytics so that more money will pour in."[101] So sadly, there are groups who have something to gain from the exploitation of the impoverished and abused. Simultaneously, these groups seek to reap the reward but downplay their responsibility[102], as we explored with the notion of 'glitch'.

We can become distracted from these harmful cycles, not only from the overinflation of technological improvement, but the "pervasiveness of race talk," which "can serve as a proxy for more far-reaching social progress."[103] Seeing some people of color in 'high places' cannot distract us from the socioeconomic disparities for a majority of their communities. We can champion the social progress of this successful subgroup, but we are concerned with those made vulnerable daily by WMD, not simply those who were able to make it out of these cycles of abuse. Companies' diversity initiatives are often poor models themselves, as they emphasize

---

[100] Benjamin, *Race After Technology,* 9.
[101] O'Neill, *Weapons of Math Destruction,* 13.
[102] Benjamin, *Race After Technology,* 76
[103] Ibid, 25.

cosmetic change over legitimate improvements. We need to bring focus to root corrective measures, not trivial improvement that would still perpetuate harmful practices if the core issues are not resolved. Technological advances cannot equate social good, even if some add to societal betterment.

As for the watchdogs of marginalized groups, throughout American history there is overwhelming evidence of the "color of surveillance."[104] From the wiretapping of Dr. Martin Luther King in the late 1960's, to the suspicion of Bolshevism among returning Black WW1 servicemen, to the "mosque crawlers" monitoring and reporting on Muslim neighborhoods post 9/11[105]. Today, this posture is manifested in places like the surveillance of Black Lives Matters activists and American immigrants.

Even without malicious intent of surveillance—as in the case of the Uyghurs—fear is still rampant among exposed communities who constantly experience the privacy breaches of oversurveillance. The vulnerable communities of the US fear the abuse of their information and bodies because they do not fit the fabricated social norms, which thus pollutes every area of their lives through generational abuse and present-day discrimination.

In a world where everyone is watched on some scale, those who come from privileged backgrounds have the notion of equal surveillance—a one size fits all privacy risk[106]—because they do not fit the demographics of the significantly over-surveilled and typically unaware of the imbalance. Thus, they have less incentive to invest in policy and regulation design to protect all, until they are made aware and/or directly affected. For example, by the time of Snowden's 2013 exposure of the NSA surveillance scandal of all citizens, the DEA and NSA had already been

[104] (Bedoya 2016)
[105] Ibid.
[106] (Madden 2019)

watching and investigating immigrants for decades without major rebuttal[107]. Likewise, while the Uyghurs are forcefully encouraged to submit their lives to this monitoring, the majority of Han Chinese are largely ignored and left to peace, implementing the idea of segregated surveillance. The Han Chinese are waved through 'green channels' such that in one city, there are "separate yet overlapping worlds on the same street".[108] If majority groups experienced the same intensity of surveillance, there would be uproar for protection of rights. This further marginalization of exposed communities makes them more vulnerable because increased surveillance essentially ensures they will find more damning data on that group, similar to predictive crime models. The privileged who are typically exempt will often declare that vulnerable communities should not care if they have nothing to hide, until it is their own rights being impeded on.

We should not wait until all groups are significantly impaired by faulty data, algorithms, and practices to take action, but use this exposition of potential harm to begin to mitigate the damages. Although complex factors like racial discrimination are extremely ingrained in our institutional structures, and thus difficult to alleviate it confounding nature in data science, we need to take the steps towards possible mitigation of harm.

CONCLUSIONS

Institutional racism and racist conditioning reinforce stereotypes and encoding them into our data only ramps up these harmful cycles further. When technology confirms our racial bias, we assume they are true and our systems intake even more data based on the biases. This cyclical exploitation weakens vulnerable and marginalized groups further. Then, it seems like models like

[107] (Bedoya 2016)
[108] (Buckley and Mozar 2019)

predatory loan algorithms, recidivism sentencing models, and credit-dependent applications are meant to keep the impoverished poor forever[109], despite any good intentions.

It can become easy to avoid topics of implicit bias, poverty abuse, social-strata, education and employment disparities, and over-policing when we are seeing the positive impacts of widespread profit, technological innovation, medical advancements, and all the benefits of a highly technological society. There certainly is progress, but it masks the digression of equity due to faulty data science. Then the protection of people and their privacy can begin to crumble under the guise of pure advancement.

Anyone deeply involved in changing the way someone lives—such as determining their parole or if they are eligible for a home loan—has a duty to examine their practices and understand the institutions they maintain. This is not meant to demonize data scientists, but instead call them to action and to awareness of their authoritative standing and impact on the progression of society. Also, to advocate for intentional diversity of teams, so to better recognize potential malfunctioning data cycles. Modernity demands challenging the systems around us that reinforce racial biases or the marginalization of any group, and data scientists have the privilege and responsibility to investigate and mediate these issues. Abuse of data simply could not proceed extensively if the internal processing was exposed and reviewed, bringing the 'black box' into the light.

[109] O'Neill, *Weapons of Math Destruction*, 155.

Appendix

Honors Research Symposium Presentation

May 21, 2022, Seattle Pacific University

In the modern day and age, data is in high demand, blooming with extreme potential, and produced constantly from every active member of society. Given its massive presence in our current climate, we are highly dependent on algorithms for our internal processing and assessments that were previously determined by man alone.

In general, the intention of these models is social betterment, such that they are designed to work more efficiently, more effectively, and are intended to weigh all data points equally as to remove potential human bias. Both the subjects and the employers of these algorithms are meant to benefit, or at least this is the image that data science puts forth. It was not long ago that a common illusion consumed our society, one that religiously perceived technological advancement as equivalent to societal improvement and thus saw the related practices, like data science, as the unparalleled authorities of knowledge, given their intricate structures and widespread integration in our everyday lives.

However, our personal data is highly interconnected and when supposedly neutral models intake all data points as equal, they will inherently ignore the socioeconomic disparities of marginalized communities, and potentially punish their vulnerable positions because of their preexisting disadvantages.

We can explore the interconnectedness of poverty and race with zip codes through the US' history of redlining. The post-abolition period saw the immediate development of state and local Jim Crow laws that barred Black citizens from opportunities like employment, well-

managed government institutions, mortgage loans, and housing prospects, so the type of areas and properties people of color qualified for and could afford were extremely restricted, which resulted in residential segregation. Additionally, since they were forced to buy or rent in lower income neighborhoods, they appreciated less on investment properties compared to White borrowers, enacting a dangerous feedback loop of generational wealth disparities.

So, even when the Fair Housing Act of 1968 was passed to prohibit racial discrimination, generations of wealth disparities already made them often unqualified for credit opportunities, and they continued to locate in the same areas typically. Thus, as this trend bleed into the following generations, they have appreciated less wealth on average such that zip codes are inherently tied to racial and wealth divisions today.

Keeping these connections in mind, when we examine algorithms such as criminal risk assessments like the LSI-R, we see that inquiries of 'basic' information in their questionnaires become racially charged due to the byproducts of generational disparities. This test is designed to gauge the level of risk an inmate can impose on their community, based on a series of 54 personal questions, such as home location, frequency of police interactions, first involvement with police enforcement, education, and financial history, all which have distinct connections to racial divides.

Frequency of police interactions and first involvement with police enforcement is highly correlated to the oversurveillance of Black communities, which can be traced back over a century ago as well, with the watching of slaves and freed Black citizens to prosecute 'black misconduct', then to the observation of supposed 'race agitators' like Martin Luther King Jr. and Malcom X during the civil rights movement. Today, predictive crime algorithms typically send police enforcement to colored neighborhoods more frequently due to the predisposition to

oversurveillance, and then often fail to consider how this trend will inherently lead police to record more crime in the areas they are increasingly watching. Thus, young black men are more likely to have earlier and more frequent run-ins with the police than their Caucasian counterparts because of this imbalance. Not to mention their increased likelihood of being charged for the same activity. So, when questionnaires demand information like an inmate's first engagement with law enforcement and frequency of interactions, they can be punished for coming from a neighborhood that is predisposed to increased surveillance instead judged according to their individual actions.

Neighborhood crime data is not always recorded ethically in the first place, such as several jurisdictions in New York state have been known to "juke the stats" with intentional data manipulation like falsifying minor crimes to meet quotas, according to authors Schultz and Crawford of the study "Dirty Data, Bad Predictions". When thirteen jurisdictions in New York were under investigation for such illegitimate police practices, their corrupt data was still being inputted into the predictive crime algorithms, creating a harmful feedback loop of skewed information. We do have to acknowledge that yes, there are areas with higher crime rates, which stems from a multitude of causes and thus do require increased police enforcement, but their residents should have the right to proper data collection regardless.

This algorithmic punishment for race and location is jarring especially considering the case of the two young Black women in Coral, Springs Florida, who played with another child's toys on the street and were charged for theft—when they were predicted to be riskier and more likely to commit another crime than a seasoned white offender of armed robbery, there is clearly something askew in this algorithm. This is no glitch, as a decent proportion of crime risk assessments across the country are being investigated for racially charged outputs.

Even a perfect model cannot separate ‘good’ data from ‘bad’ data such as with our case of juked statistics. Mathematician Cathy O’Neill describes this as the “garbage in, garbage out” concept. Faulty data will continue to instigate a system that produces inequity because the data models are solidifying illegitimately collected data as fact.

A model that embodies this potential for good but is skewed by poor data collection and implementation is the IMPACT teacher evaluations of the District of Colombia. This assessment of classes’ math and language proficiency is designed to celebrate high-performing educators and reform underperforming classrooms. When the previous year’s scores were embellished to look more impressive, then following year, students’ progress looks stagnant or declining, and the next teacher was punished for it. Furthermore, if this assessment is applied as the deciding factor of employment status, despite personal testimonies or taking account for any confounding impacts on student performance, it becomes harmful and inaccurate. As a supplemental tool, the IMPACT scores provide helpful insights, but a model alone cannot analyze external factors that are vital considerations for education quality levels. The score alone should not and cannot fully embody the reality of an educator’s career.

Looking at the process of reviewing data for accuracy, we see that it is increasingly tedious and complex as it requires assessments of consistency, formatting, plausibility, quality, handling, units of measurement, quality of collection method, and even more. Additionally, this process is informal and not accurately enforced as companies rely heavily on self-regulation. Most data scientists champion data sharing because collection is expensive among all resources, and their self-regulation allows them to. Truly, open data sets have immense potential for good as data scientists can work more efficiently, but if faulty data is circulated between researchers and companies, it can lead to more defective data outputs. Especially if skewed data is not

scheduled to expire, it can continue to harm subjects across many arenas for an extended period of their lives. Such as California's gang database, which was created through a combination of zip codes and racialized names. Thus, it was majorly composed of Blacks and Latinos, even including babies younger than a year old. And once a citizen is categorized, it is hard to alter one's standing as this database was not scheduled for liquidation for more than 100 years, despite its racially charged methods of collection and obvious inaccuracy. However thankfully today California has enacted the Consumer Protection Act which allows residents to ask companies to delete data or prevent the sale of their data. Likewise, the European Union has passed the General Data Privacy Regulation which enforces transparent data processing, minimizing data collection, maintaining accurate data, accountability of data controllers, time limits of data storage, etc. The US has yet to see a nationwide law like this as companies simply promise ethical data practices while maintaining their "black box" of data analysis. This black box essentially embodies the opaqueness of data processing and practices. This signifies how our data goes in and results come out of algorithms, yet all the while the inmate is shielded from the questionnaire's procedures, the educator does not know why they are being fired, and the borrower has a seemingly random spending limit decrease.

American Express has been known to cut spending limits of customers who shop at particular 'lower-income' stores, such as Walmart, thus lowering their credit and driving up borrowing costs. Likewise, Capital One created a scoring system to act as a proxy for credit scores, based on online shopping history to gauge prospective borrowers' wealth so to create corresponding advertisements. Those who score low on this scale receive advertisements for credit cards with lower credit limits and higher interest rates, enforcing another malicious feedback loop for those of within the lower income demographic. The categorizing of buyers

into similar socioeconomic and educational buckets may coerce them into the same actions simply from the advertisements they receive. These algorithms embody Princeton professor Ruha Benjamin's question of "what do 'free will' and 'autonomy' mean in a world in which algorithms are tracking, predicting, and persuading us at every turn?" For the authoritative positions like the judges and lenders, the difficult decisions are made for them, which only reinforces this inflated trust in the mathematics of data science. At the same time, the everyday citizens are being abused and still may be completely ignorant to the harm.

On this thread, we cannot continue to blindly support colorblind models as our lone assessors, as they continue these vicious cycles of harm due to the racism and implicit bias that have infected our institutions. There is growing advocacy for racially literate models, because even though models themselves are neutral to the information they are given, when the infrastructure is corrupted, the whole system is can be compromised so that equality does not ensure equity. This concept of racially literate models is unfortunately complicated considering the complexity of institutional racism, but we can start by guiding our decision making processes away from such heavily algorithm dependent structures. Human agency should be maintained in some form, so that there is an opportunity to appeal to human rational and situational awareness, as unordinary circumstances cannot be accounted for by algorithms. Within this, there is also a call for racially literate data scientists, so that they know how to examine how preexisting bias will interact with their design and if their teams have enough diverse perspectives in the first place, as to recognize potentially racially skewed elements.

The 'glitch' no longer exists as a scapegoat in an algorithm that is created and controlled by Man. Data is the epitome of interdisciplinary arenas, and thus, the technical experts need to be held ethically accountable for their models that impact lives tremendously.

We can see the many routes in which data science can potential to go wrong, within the stages of data collection, analysis, and practice. Beyond these controllable factors of data structures, institutional racism complicates this discipline on another scale altogether. For experts in the field of ethnic and racial studies, like Black Studies professor Christina Sharpe, they see the current 'climate' or 'weather' of our society as anti-Black. And while individual hate may be more sporadic and be socially rejected, institutional indifference determines the fluidity of social structures and is harder to alter significantly.

Modernity encompasses the dichotomy of valuable and devalued vulnerable bodies. The information of these communities, which are typically made up of people of color, is highly valued and overexposed, while their bodies and their persons are devalued and not truly perceived. This evokes the idea of "those who are watched but not seen". People are quantified as data points and this intellectual property is highly valuable. So, the tailored algorithms of companies are heavily defended and obscured from the public eye as to ensure the preservation of their practices. As design is optimized, racial bias slips into the seams of algorithms and thus forms the New Jim Code coined by Ruha Bejamin, which stratifies people into stigmatized buckets because of an unfaltering confidence of colorblind models and the reality of data manipulation. Thus, people can become expendable, as long as their data is retained for profit.

There certainly are societal improvements that come from the field of data science, but these can mask the digression of equity that is happening alongside areas of progress. The protection of people and their privacy can begin to crumble under the guise of pure advancement. We cannot allow this harm to continue, so we must expose the internal processing of data science and demand accountability of their authoritative positions.