DATA IS NOT NEUTRAL
Charlotte McClintock
transcript of a talk I gave to students at an honors research symposium

Algorithms built on math and statistics now govern our lives, from determining what news we see, to predicting what we might buy, to making decisions about where to send police officers. I want to convince you that understanding data and algorithms is vital to be an informed citizen in the coming decades. I want to empower you to engage in this field and make you believe that your voice is valuable and desired in this discussion.

I. DEFINITIONS

First, I think it's helpful to set some definitions for this conversation. Shared language and understanding is so important to being able to engage in conversation. Data science as a field is filled with redundant confusing jargon, with often different word choices for every subfield. Data science itself is primarily a rebranding of statistics with some new methodology enabled by greater computing power.

What do we mean by *data*?
All information is data. Data can be quantitative and numerical, but it doesn't have to be. Photos, videos, written and spoken language are all now considered valuable data. We use data all the time to make decisions in our daily lives. Reviews for clothing or restaurants give us data. Checking grade distributions for classes on VAGrades or CourseForum while picking classes? You're using data.

What do we mean by *algorithms*?
An algorithm is just a process or a set of rules for setting priorities and making decisions. Algorithms can be simple or complex. We often think of mathematical or statistical algorithms, but any decision-making process can be considered an algorithm under a broad definition. Say we sorting umbrellas on Amazon by the number of 5-star reviews, and limit our search to umbrellas under $20 with more than 20 reviews in order to decide what to buy. That's an algorithm! Machine learning is usually used to describe predictive algorithms.

II. BIAS IN THE TECHNOLOGY FROM DATA

So often, we think of data and math as neutral. When math is first taught to us, we learn it in the abstract. Algebra and calculus contain no systemic biases or societal privilege, they simply give us a mechanism for quantification and computation. But the application of math to the social and political realm has brought all of the opinions and debates of culture and politics into the quantitative realm. Data is created by governments, NGOs, and companies and the numbers lose the neutrality of abstract integers when created by a system to serve a purpose. Algorithms built using biased data can often pick up the biases as meaningful trends within a system, when they're really just replicating systemic bias.

Let's consider an example that was reported recently by Quartz. Algorithms are increasingly used in hiring to screen applicants and decide which applications to read. An employment firm was approached by a startup with an algorithm which screened applications, read resumes, and predicted job performance. The algorithm was intended to simplify, improve, and accelerate the hiring process. What would you expect might predict job performance from a resume and application? (GPA? Number of extracurricular involvements?)

The firm audited the algorithm found two data points were strongly associated with high job performance scores: if your name was Jared, and if you played high school lacrosse.

The algorithm was built on training data that was biased. Authority figures in the workplace are predominantly white and male, and assess performance for young workers like themselves more highly. The algorithm here was also overfitted to the data and picked up statistical noise. This example is amusing and relatively harmless; the bias was identified before it was widely deployed.

The Quartz article also mentions another attempt by Amazon to create a similar algorithm for hiring. After the company trained the algorithm on 10 years of its own hiring data, the algorithm reportedly became biased against female applicants. The word "women" would cause the algorithm to specifically rank applicants lower. Amazon killed the project, but hiring is a time intensive process for companies and these algorithms are only going to become more common.

In an attempt to quantify behavior and capacity more and more to reduce human labor, we should expect algorithmic bias to become more dangerous and applied on a larger scale. Examples of bias are everywhere. Image classifiers frequently classify black women as men. Virtual reality headsets built for and tested on men made women nauseous. Risk score assessments used to advise judges in decisions about pre-trial release mislabeled as high risk more often than white defendants. Online ads show men jobs that pay better.

Bias in our data created by existing disparities creates biased technology, which in turn perpetuates existing disparities if not caught. No data is neutral. All attempts at simplification, reduction, and quantification of information require some judgement on what is most important, and how things should be recorded and grouped. We need to acknowledge the opinions inherent in our information and data and treat the results of analysis accordingly.

III. BIAS IN THE PRACTICAL APPLICATION OF TECHNOLOGY

"Solving" bias in technology is important but can lead us to overlook another significant issue: the potential for application of even unbiased technologies to practical problems in a biased manner.

Technology can often be harmful by enabling precise targeting or harsh bright line rules which are unforgiving to human concerns passing through the system.

In "Automating Inequality", Professor Virginia Eubanks details the transition of the Indiana social services system to automation. The transition was made in the name of efficiency, reducing fraud and waste in the distribution of welfare. The new system was built to be unforgiving. Rather than scanning in previously existing documents, the new system required every household to re-certify their eligibility in a short window of transition. The new system removed the caseworker role and replaced them with call centers where people could call and receive one-off assistance, but had no consistent resource to go to. The system was set up so that any small error, a missing date or initials, resulted in a "Failure to Cooperate" response where the services were terminated without informing the individual of the specific missing information. This resulted in thousands losing life-saving Medicaid coverage, nutritional support, or other vital services.

We saw another example in 2018 in Georgia. The Secretary of State's office under Brian Kemp decided to crackdown on potential voter fraud with an "exact match" system. The office invalidated about 53,000 voter registrations where the recorded name differed at all from other state databases, anything from one missing letter to a missing hyphen in the resident s last name. An analysis from the Associated Press found that while 32% of the Georgia population is Black, around 70% of flagged registrations were from Black residents.

Even if we could root out bias from our algorithms, we need to consider how the technology will be applied and which communities will be affected. How do we reconcile a desire for better facial recognition algorithms for people of color with knowledge that the algorithms we build may be used against them in surveillance and policing? When we know technologies are going to enable increased surveillance of vulnerable people, are they safe just because they're unbiased in regards to race or gender? Even theoretically neutral technology can be applied to systematically disadvantage marginalized groups, and our conversations about bias in technology need to extend to applications of the technology in biased ways.

IV. WHERE WE GO FROM HERE

We are members of the Information Age. Data is increasingly valuable as computing power increases to enable more powerful analysis than ever. Capitalism is shifting to a new market based on selling information, predicting, and shaping human behavior. Too often, data practitioners have little to no experience in the context of the analysis and think of this as an asset that preserves their neutrality. No data exists independent from the systems and actors which created it, and no analysis is truly objective. We project our experiences and our biases onto our mental context for the data, and make decisions accordingly, from choosing controls for a regression analysis to deciding how to categorize race, ethnicity, and gender to deciding when a result is substantively significant.

Two quotes sum up this sentiment very well. The first is from Jake Porway, the CEO of DataKind: "AI serves the systems that already exist and we have to be intentional to change that." The second is from Professor Julia Powles, "Addressing bias as a computational problem obscures its root causes. Bias is a social problem, and seeking to solve it within the logic of automation is always going to be inadequate."

Data used well has the power to do so much good. Narratives and anecdotes are useful in persuasion but can be deeply misleading when used to characterize an entire population or set of experiences. Data and quantitative analysis are incredible tools to inform policy and political decisions and improve outcomes for underserved and marginalized citizens who aren't being well served by existing services. Looking into the trends in data can challenge our preconceptions and expectations, show us unexpected insights, and help us find where vulnerable people might be falling through the cracks of our systems. Quantitative analysis can provide important insights to the philosophical debates on public policy, politics, and social impact work.

In the future use of data, we need to walk the line balancing data misuse with data missed use. We need to acknowledge the biases data contains, and the opinions we bring to the context of an analysis.

Let's democratize data science and analysis. Data science is fundamentally an attempt to find meaningful patterns in information. I believe those who use data should have experience and context in the subject area when going into an analysis, listening to the individuals they hope to serve and attempting to gain a deep understanding of the context of their work. I'm in these spaces all the time, and we really need people like you bringing your content and domain knowledge. So many people making critical choices that will shape our future are agnostic to the domains you are immersed in. We have to include women, people of color, LGBT folks, and low-income people in decisions that affect them. We need transparency and accountability and inclusion.

I'd like to leave you with questions rather than answers:

- What does this technological future mean for us as human beings?
- How will this change the future of work, the future of interaction, and the future of surveillance?
- How can we better protect privacy and ensure transparency?
- What's *your* role in this space?