

Ethics and Data Science: Where to Start?



Abeba Birhane

Cognitive Science PhD
candidate,
School of Computer Science,
UCD

 @Abebab



Siobhán Grayson

Data Science PhD candidate,
Insight Centre for Data Analytics
& School of Computer Science,
UCD

 @siobhan_grayson

Outline

Part 1: Why are we here today talking about ethics at data science meetup?

- Are data and automated systems neutral?
- Examples

Part 2: Towards thinking ethically - open problems

- How can we incorporate ethical approaches into our work?
- What issues require further thought and discussion?

Common Assumptions Data & Automated Systems



- Objective
- Neutral
- Straightforwardly represent reality

Data - misconceptions and (incorrect) assumptions

- Objective
- Neutral
- Straightforwardly represent reality

Data in reality

- As subjective as humans recording it
- Political
- Messy
- Often incomplete
- Sometimes fake
- Full of complex human meanings

“Data and data sets are not objective; they are creations of human design. We give numbers their voice, draw inferences from them, and define their meaning through our interpretations. Hidden biases in both the collection and analysis stages present considerable risks, and are as important to the big-data equation as the numbers themselves.”

(Crawford, 2013)

Data often reflect existing societal inequalities,
norms and practices

BIAS IN



BIAS OUT

People make important choices

- Which data to use and by default which to exclude
- How to weigh/analyse it
- How data are manipulated & reported
- How good is good enough for an algorithm to perform

And people are biased...

Bias

A CV with a white-sounding names will receive a different response than the same CV with a black-sounding name. (The National Bureau of Economic Research)

Women's chance of being hired in symphony orchestras increases between 30% and 55% in blind auditions. Goldin & Rouse (2000)

Cognitive biases can skew results – they pose a challenge to data science

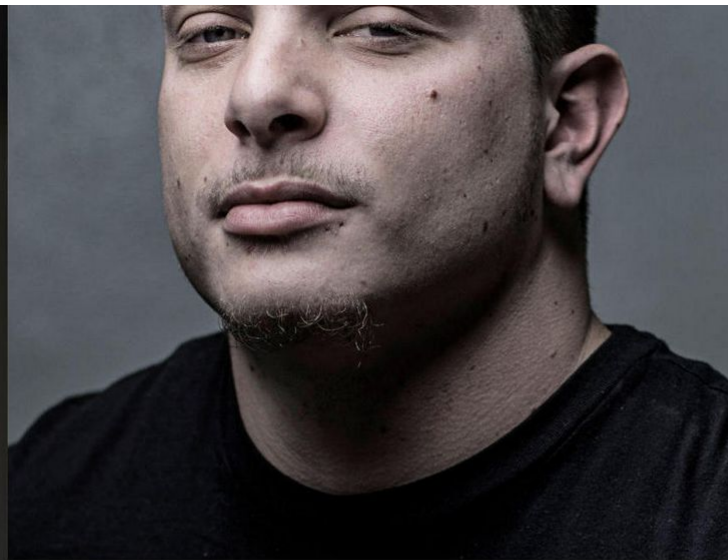
We are often inclined to try to solve these problems stemming from human bias by turning the decisions over to machines

Machine Bias

In theory, more and more decisions increasingly handled by algorithms should mean that human biases and prejudices should be eliminated

Algorithms are, after all, “neutral” and “objective”

They apply the same rules to everybody regardless of race, gender, ethnicity or ability



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica



Machine Bias

Mathematical models give the illusion of “neutrality” O’Neil (2017)

Automated systems perpetuate stereotypes as they learn from real-world data

They pick up damaging biases that have long plagued human culture



Alex Shams

@seyyedreza

Follow



Turkish is a gender neutral language. There is no "he" or "she" - everything is just "o". But look what happens when Google translates to English. Thread:

o bir aşçı
o bir mühendis
o bir doktor
o bir hemşire
o bir temizlikçi
o bir polis
o bir asker
o bir öğretmen
o bir sekreter

o bir arkadaş
o bir sevgili

onu sevmiyor
onu seviyor

onu görüyor
onu göremiyor

o onu kucaklıyor
o onu kucaklamıyor

o evli
o bekar

she is a cook
he is an engineer
he is a doctor
she is a nurse
he is a cleaner
He-she is a police
he is a soldier
She's a teacher
he is a secretary

he is a friend
she is a lover

she does not like her
she loves him

she sees it
he can not see him

she is embracing her
he does not embrace it

she is married
he is single

The decisions delivered from automated systems may not matter if these systems are recommending what books we might want to buy next based on our previous purchase

However, the stakes are higher if the system is diagnosing illness, or holding sway over a person's job or prison sentence

Understanding the legal, cultural, historical, and social contexts and the nuanced, complex and fuzzy boundaries are crucial

Not a day goes by without tech companies making headlines



Skyscrapers



Airplanes



Cars



Bikes



Gorillas



Graduation



Women less likely to be shown ads for high-paid jobs on Google, study shows

Automated testing and analysis of company's advertising system reveals male job seekers are shown far more adverts for high-paying executive jobs

Hiring Algorithms Are Not Neutral

by **Gideon Mann** and **Cathy O'Neil**

DECEMBER 09, 2016

 **SAVE**  **SHARE**  **COMMENT**  **HH** **TEXT SIZE**  **PRINT** **\$8.95** **BUY COPIES**





AGE DISCRIMINATION



New Allegations Added to Lawsuit on How Facebook's Targeting Tools Helped Advertisers Exclude Older Workers

A federal suit filed in December claimed older workers missed out on job opportunities because ads on Facebook targeted younger users. Now plaintiffs say Facebook's tools and algorithm gave employers ways to intensify the effects of such targeting.

by Peter Gosselin, May 30, 2:03 p.m. EDT

Tech's Ethical 'Dark Side': Harvard, Stanford and Others Want to Address It

[Leer en español](#)

By NATASHA SINGER FEB. 12, 2018



Laura Norén, who teaches a data science ethics course at New York University, said, "You can patch the software, but you can't patch a person if you, you know, damage someone's reputation."

Sam Hodgson for The New York Times

RELATED COVERAGE



Should Your Driverless Car Hit a Pedestrian to Save Your Life? JUNE 23, 2016



How Tech Giants Are Devising Realities for Artificial Intelligence SEPT. 1, 2016



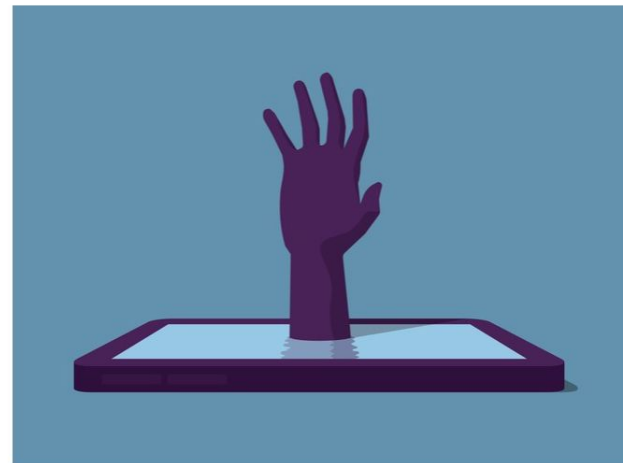
STATE OF THE ART
Facebook's Bias Is Built-In, and Bea Watching MAY 11, 2016



TECHNOPHOBIA
Can't Put Down Your Device? That's the Problem, Designers Say DEC. 5, 2015

ISSIE LAPOWSKY BUSINESS 02.08.18 10:52 AM

ETHICAL TECH WILL REQUIRE A GRASSROOTS REVOLUTION



The Center for Humane Technology wants to liberate us from tech addiction—and that starts with the people, not companies or Congress. GETTY IMAGES



“I’m just an engineer”

(February 2018 Researchers from UCLA presented a paper on predictive policing)

Ethical questions inevitably arise with innovation. But they are often an afterthought

Ethics is an integral part of medical training – physicians are upheld to specific ethical standards through the practice of swearing **Hippocratic Oath**

Data scientists, like physicians and, often work to solve society's ills, hence the need to consider how society operates and how our data shape/influence it

Should data scientists be held to the same standard as physicians?

Towards thinking ethically...

Question 1

Who bears responsibility when a system is biased/unfair or used by third parties for malicious purposes? The individual who developed it? The organization? The third parties?

Responsibility - Accountability - Transparency

Ask how and why was a certain decision made?

Learn about the possible shortcoming/limitations of the algorithmic system you're developing

Compare algorithmic systems to the law

What is “good”?

How is it decided?

By who? For who?

What are we optimizing for?

Do the motivations and incentives align with the ‘justice systems’? (Predictive policing and recidivism risk assessment)

Towards thinking ethically

Representation and diversity matter

Both in your dataset and your team/perspectives

It's never going to be a solved problem - continuous revision, learning and updates crucial

Ethical ML Takeaways

- Doing “nothing” assumes prejudice and unfair treatment is a valid action
- We need better data
 - Diverse data which better reflects the real world
 - Stop using datasets which are non-representative
- We need built-in ethics-driven evaluation criteria
 - Scikit-learn disparate impact?
 - Scikit-learn equal odds / opportunity?
- You can contribute
 - open-source your work and datasets
 - volunteer with the Algorithmic Justice League or local organization



Gravado no

QCon
SÃO PAULO

Trazido a você por

InfoQ
BRASIL

Ethical Machine Learning - Katharine Jarmul

https://www.youtube.com/watch?time_continue=3&v=Y0KMK1tfQ3A



SUPPORT INCLUSIVE TECHNOLOGY

Highlight Bias

MEDIA

Help raise awareness about existing bias in coded systems

DONATE

Identify Bias

RESEARCH

Support the development of tools for checking bias in existing data and software

COLLABORATE

Bust Bias

PARTICIPATE

Stay informed about ways to help test software for bias and create inclusive data

VOLUNTEER

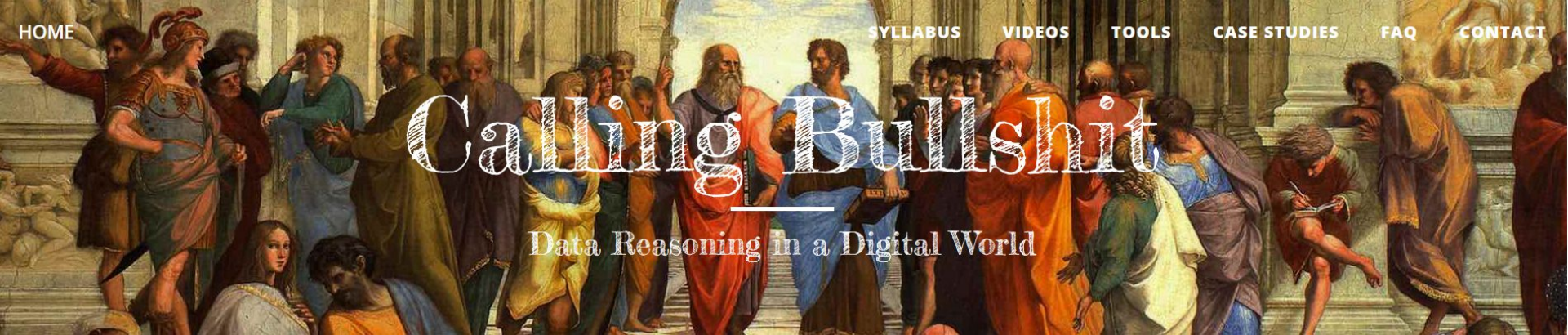
[Algorithmic Justice League](#) - run by Joy Buolamwini. Report algorithm bias, participate in testing software for inclusive training set, or simply donate and contribute raising awareness about existing bias in coded systems.

[Home](#) > [About](#)

About

Auditing Algorithms: Adding Accountability to Automated Authority is a group of events designed to produce a **white paper** that will help to define and develop the emerging research community for “algorithm auditing.” Algorithmic Auditing is a research design that has shown promise in diagnosing the unwanted consequences of algorithmic systems.

[Auditing Algorithms](#): a useful website for those teaching/interested in accountability in automated systems



[Calling Bullshit](#) offers various resources and tools for spotting and calling bullshit with relation to data, models and more

fast.ai

Making neural nets
uncool again

Home

About

Our MOOC

Posts by Topic

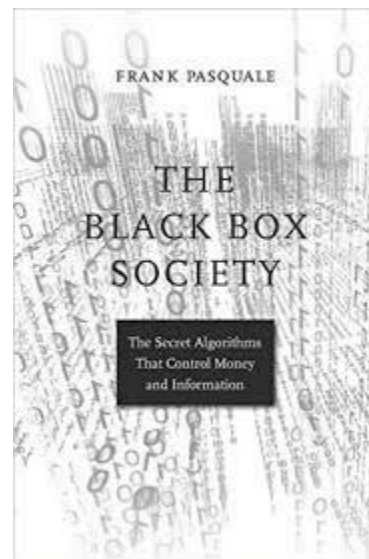
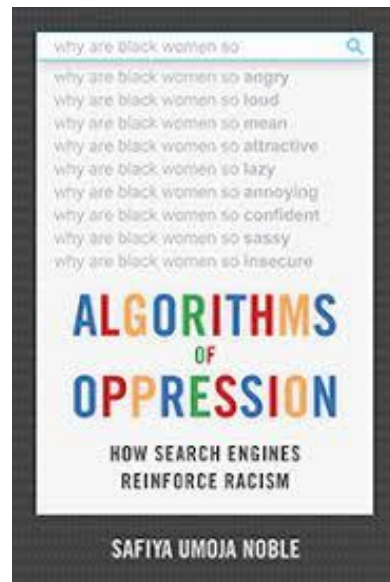
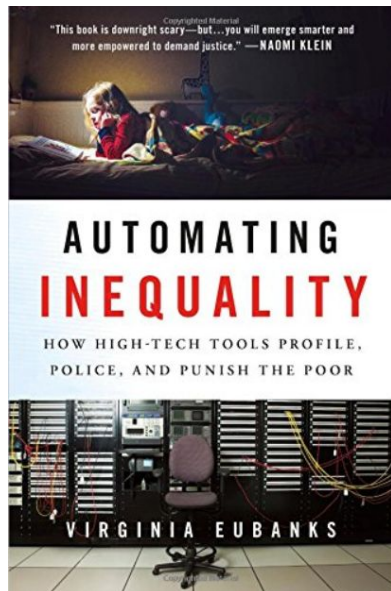
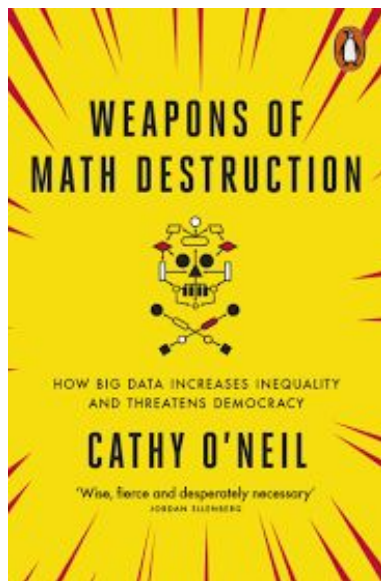
[fast.ai](#) a project that aims to increase diversity in the field of deep learning and make deep learning accessible and inclusive to all



Fairness, Accountability, and Transparency in Machine Learning

[FAT/ML](#) is a website on Fairness, Accountability, and Transparency in Machine Learning with plenty of resources and events, run by a community of researchers

Books



Resources on automated systems and bias

<https://abebabirhane.wordpress.com/>