

Responsible Data Science

Introduction and Overview

Prof. Julia Stoyanovich

Center for Data Science &
Computer Science and Engineering
New York University



NYU

TANDON SCHOOL
OF ENGINEERING



NYU

Center for
Data Science

r/ai



course logistics

Instructor: Julia Stoyanovich

Assoc. Prof. of Data Science, Computer Science & Engineering
Director of the Center for Responsible AI (R/AI)
New York University

Ph.D. in CS from Columbia University
B.S. in CS & Math from UMass Amherst



[@stoyanoj](#)

Research: data and knowledge management (“databases”)

- Responsible Data Science (RDS)
- Preferences and Voting (DB + COMSOC)
- Evolving graphs (Big Data / Systems)

And also:

- Tech policy: NYC ADS task force, NYC algorithmic hiring bill, IEEE AI policy committee, LCO of Ontario, SmartCityPhl task force
- Education & public engagement: Responsible Data Science, We are AI

Office hours: Mondays 11am-noon EST and by appointment

Course Staff

Section Leader: Aniruddha (Ani) Chauhan

Office hours: Wednesdays 2–3pm and by appointment



Grader: Jack Epstein

Office hours: Thursdays 11am-noon and by appointment

Assignments and grading

Grading: homeworks - $10\% \times 3 = 30\%$
project - 30%
final exam - 30%
labs - 10%

No credit for late homeworks. 2 late days over the term, no questions asked. If a homework is submitted late — a day is used in full.

Assignment schedule posted to Bright Space (under Course information), subject to change.

Where to find information

Website: <https://dataresponsibly.github.io/rds/> slides, reading, labs

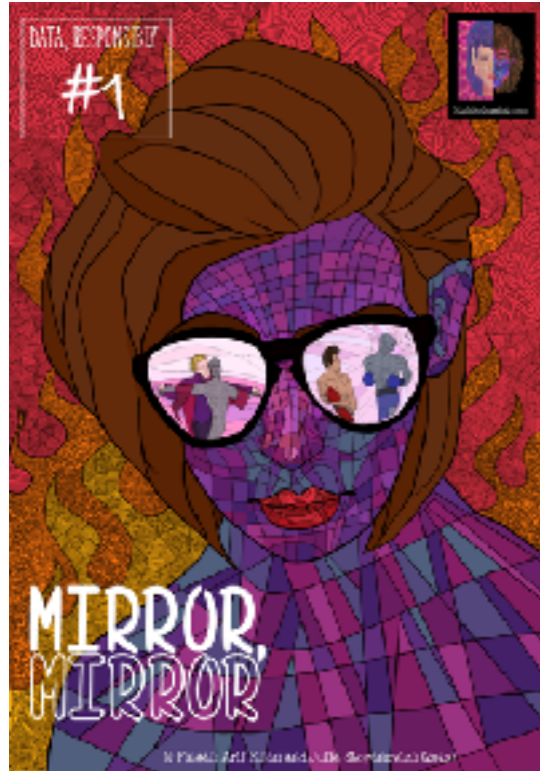
The screenshot shows the 'Fairness' page on the website. The navigation bar includes 'FAIRNESS', 'DATA SCIENCE LIFECYCLE', 'DATA PROTECTION', and 'TRANSPARENCY AND INTERPRETABILITY'. The sidebar on the left has a vertical menu with 'WEEK 1' (active), 'WEEK 2', 'WEEK 3', and 'WEEK 4'. Below the sidebar, it says 'Next module: DATA SCIENCE LIFECYCLE'. The main content area is titled 'Fairness' and contains the following sections:

- Lecture:** Introduction: What is Responsible Data Science?
 - DS-UA 202: Slides coming soon.
 - DS-GA 1017: 1 intro slides
- Topics:**
 - Course outline
 - Aspects of responsibility in data science through recent examples
 - The importance of a socio-technical perspective: stakeholders and trade-offs
- Reading:** See [Introduction and Algorithmic Fairness \(Part 1\)](#) (highlighted with a red circle)
- Lab:** ProPublica's Machine Bias
 - Colab Notebook

At the bottom right, it says 'Next submodule: WEEK 2'.

Bright Space: everything assignment-related, Zoom links for lectures and labs, announcements. **Piazza:** discussion board.

This week's reading

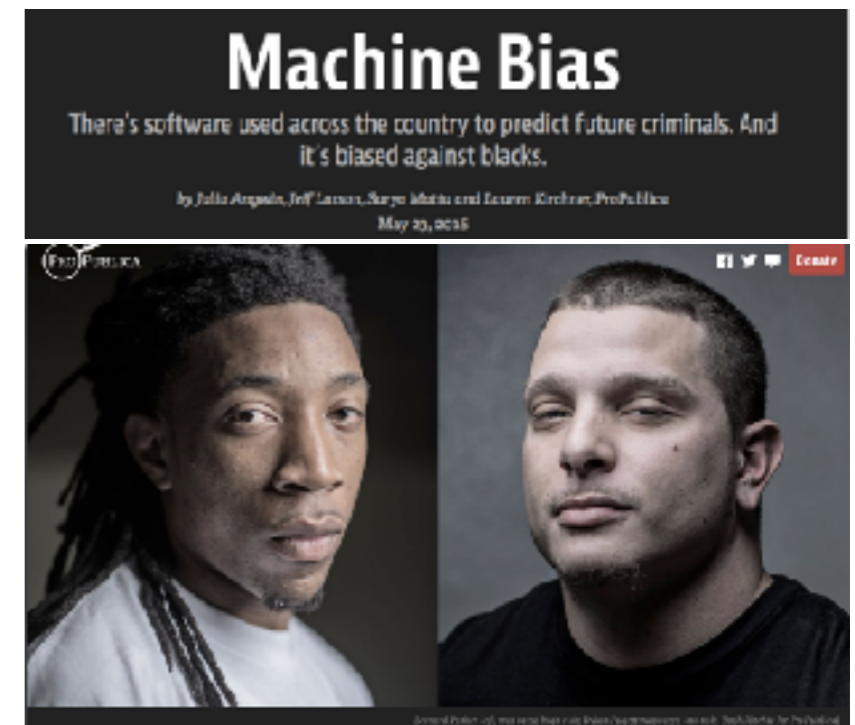


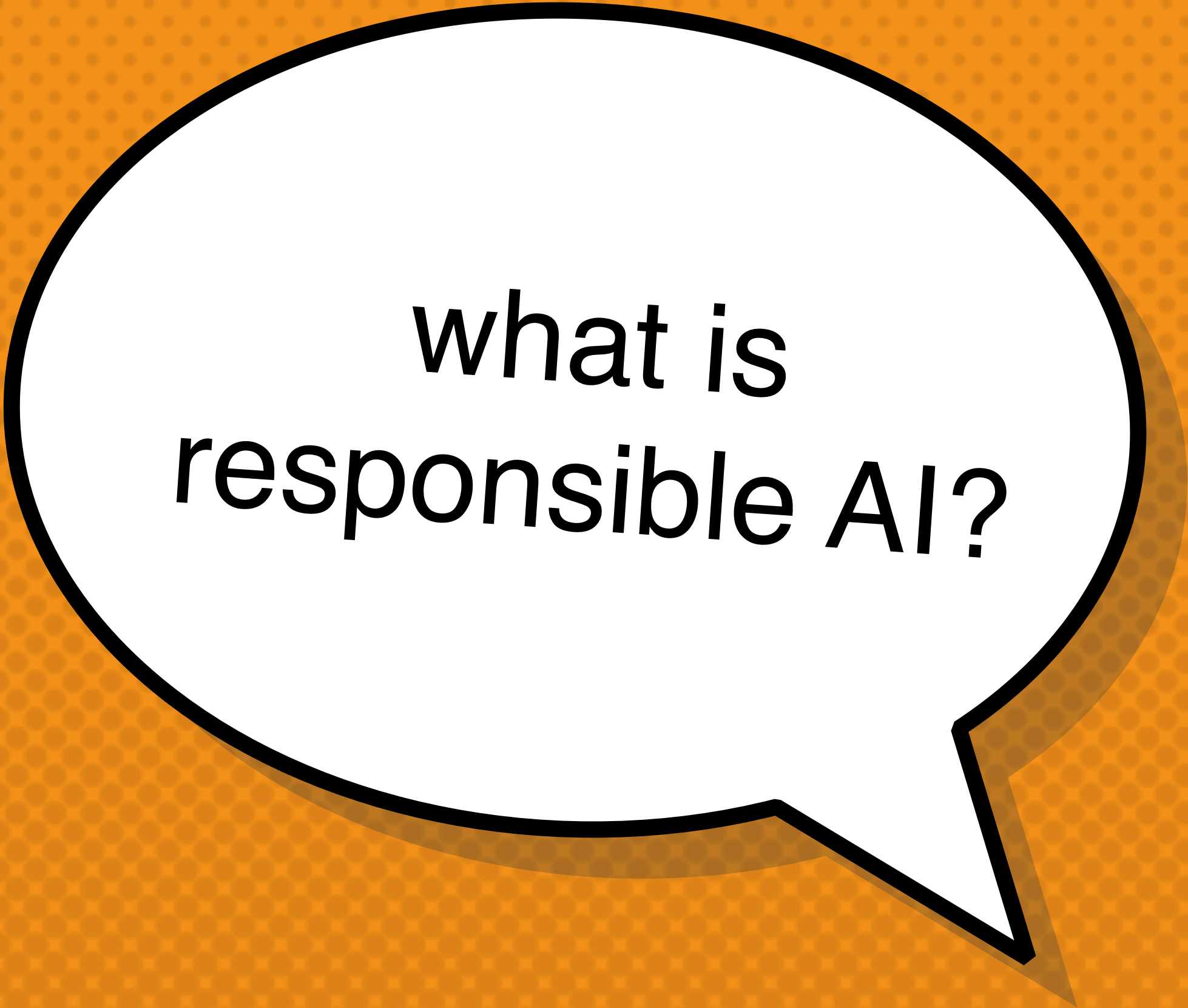
DOI:10.1145/3376898

A group of industry, academic, and government experts convene in Philadelphia to explore the roots of algorithmic bias.

BY ALEXANDRA CHDULDECHOVA AND AARON ROTH

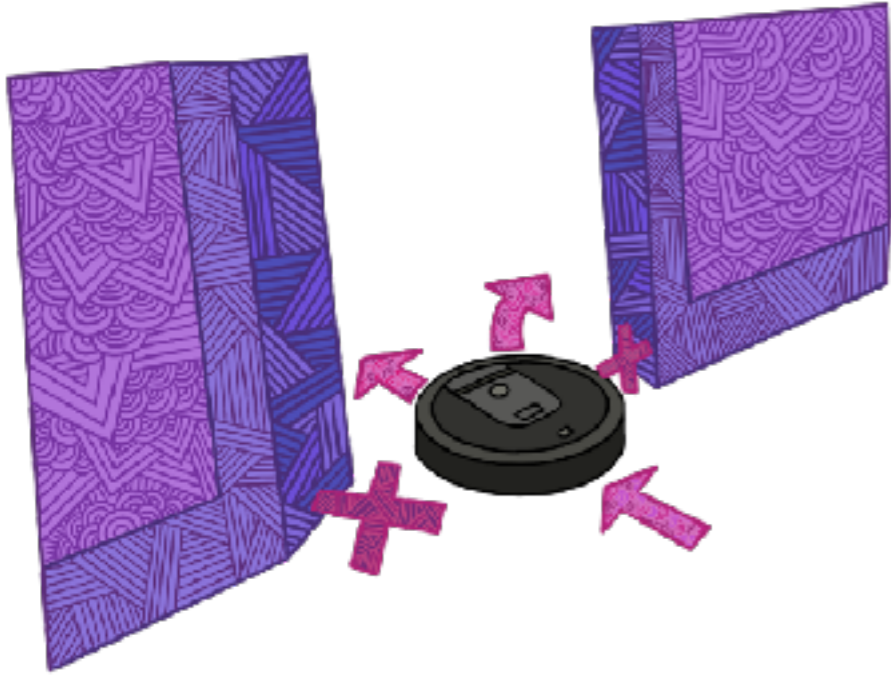
A Snapshot of the Frontiers of Fairness in Machine Learning





*what is
responsible AI?*

AI: algorithms, data, decisions



Artificial Intelligence (AI)

a **system** in which **algorithms** use **data** and make **decisions** on our behalf, or help us make decisions



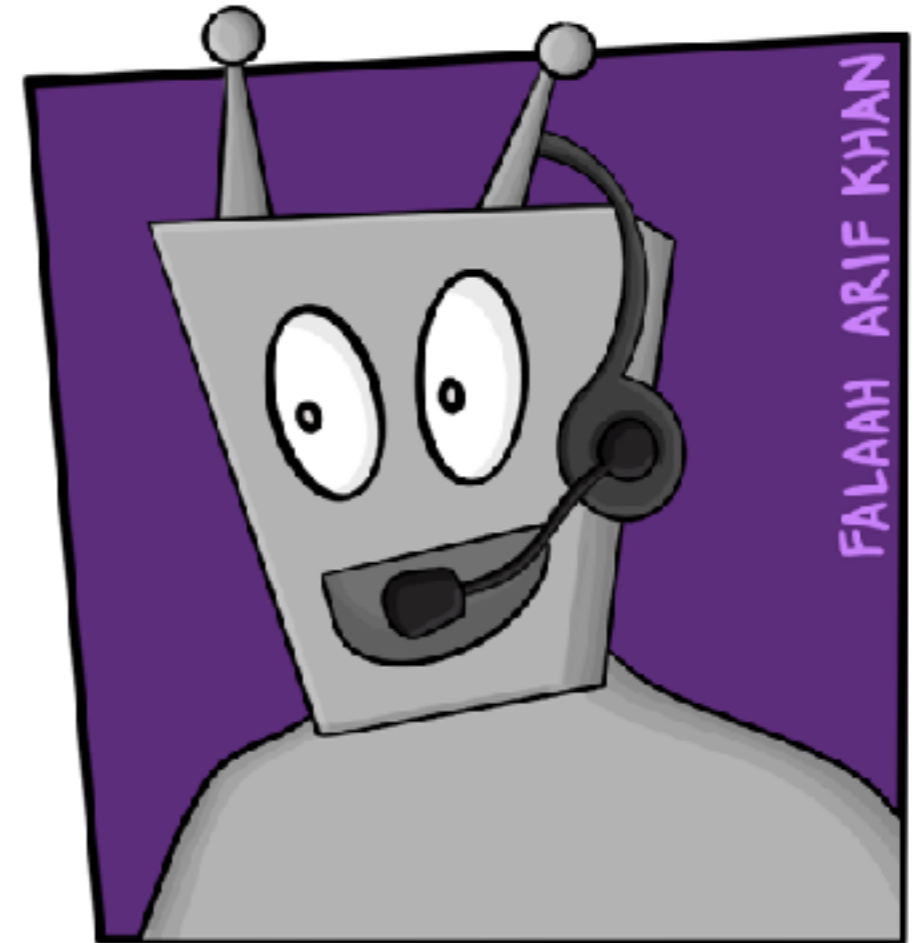
The promise of AI

Opportunity

make our lives convenient
accelerate science
boost innovation
transform government



Machines make mistakes



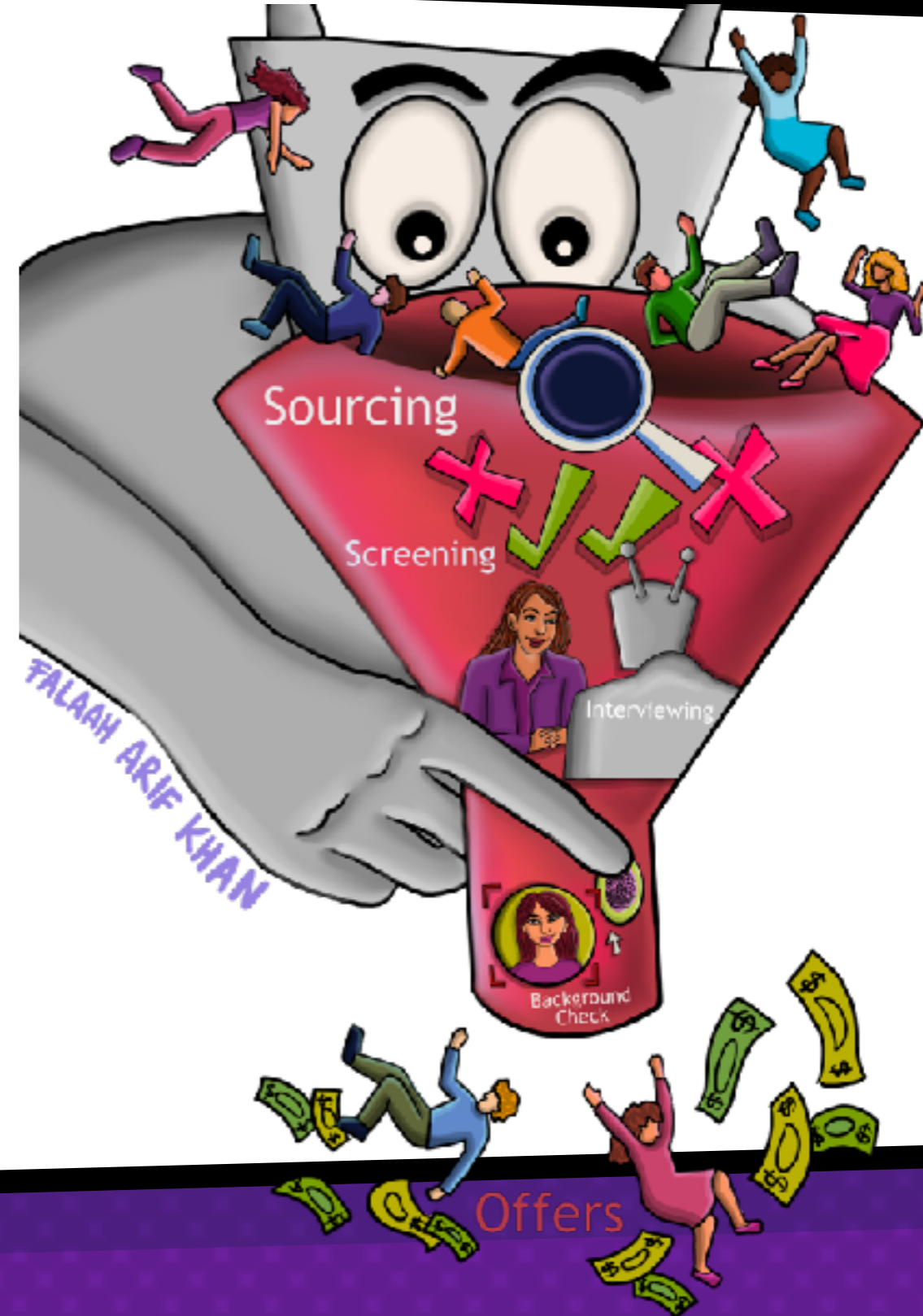
Mistakes lead to harms

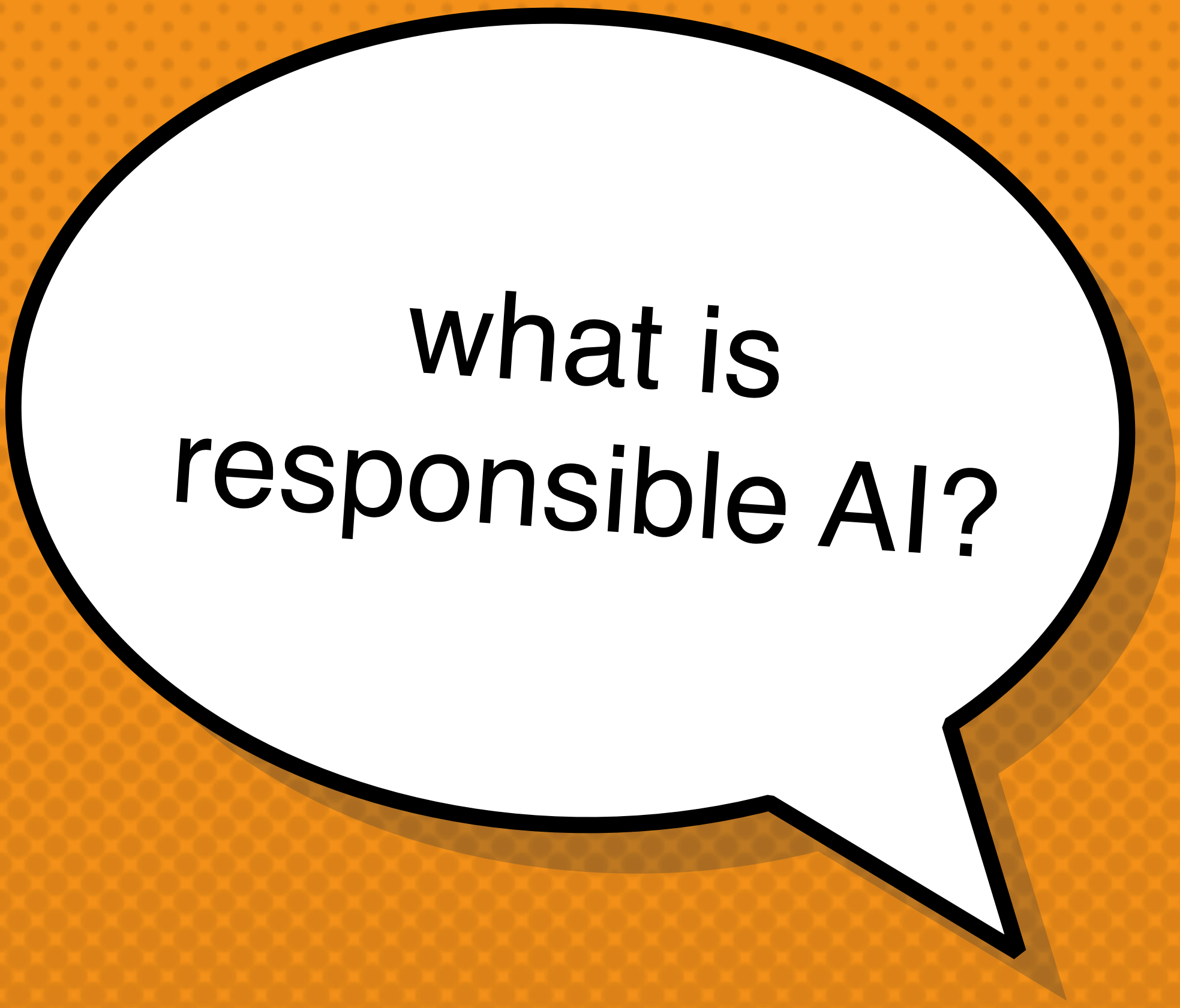


FALAAH ARIF KHAN



Harms can be cumulative





*what is
responsible AI?*



more examples

Medical imaging

FACEBOOK AI



fastMRI

Accelerating MR Imaging

What is fastMRI?

fastMRI is a collaborative research effort between Facebook AI Research and NYU Langone Health. The aim is to investigate the use of AI to make MRI scans up to 10 times faster.

By producing accurate images from under-sampled data, AI image reconstruction has the potential to improve the patient's experience and to make MRIs accessible for more people.

Positive factors

- clear need for improvement
- can validate predictions
- technical readiness
- decision-maker readiness

raw data and image datasets are available in a public repository, which contains baseline reconstruction models and PyTorch data loaders for the fastMRI dataset.

<https://fastmri.org/>

Automated hiring systems

**MIT
Technology Review** February 2013

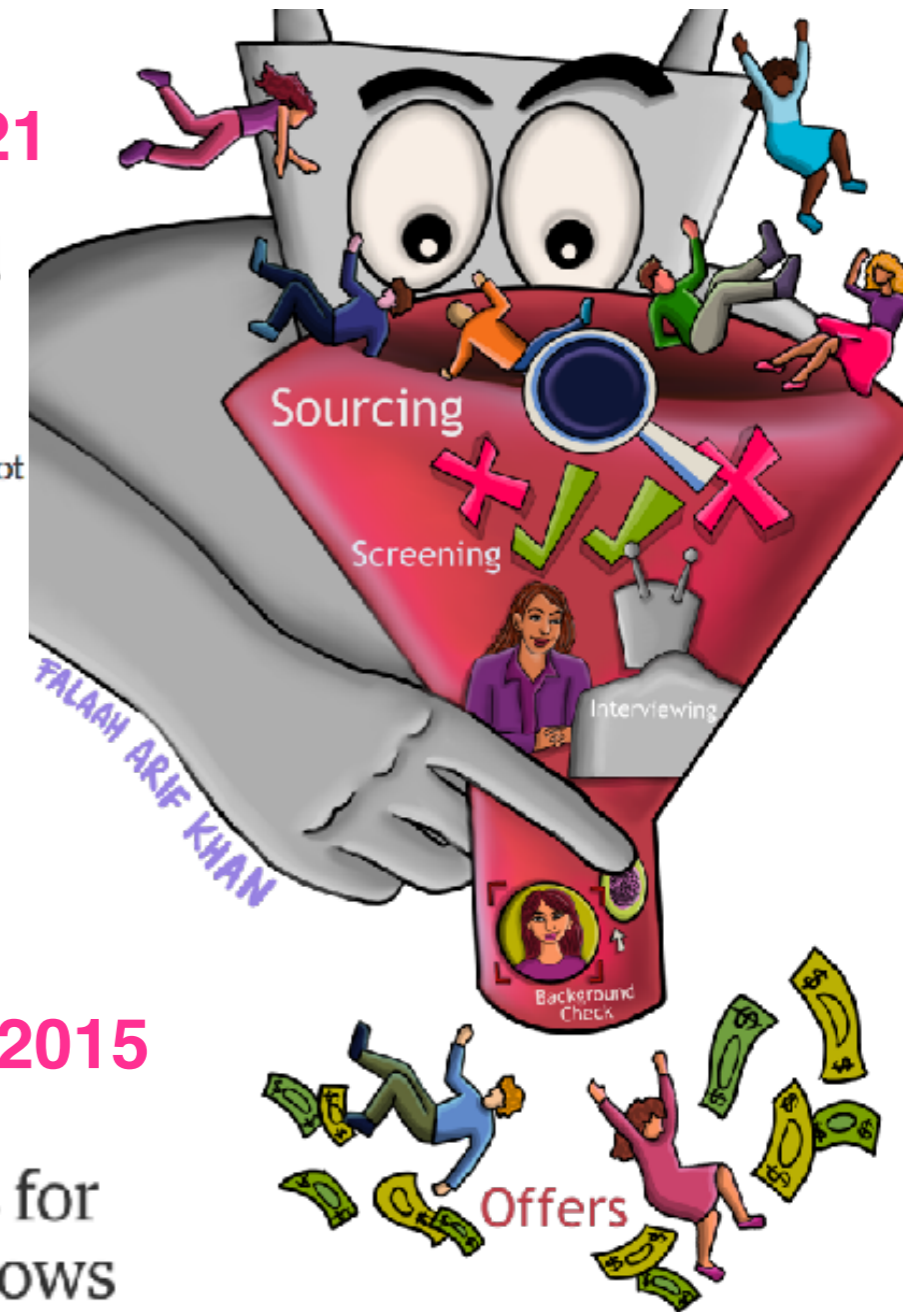
**Racism is Poisoning
Online Ad Delivery, Says
Harvard Professor**

The New York Times

March 2021

**We Need Laws to Take On Racism
and Sexism in Hiring Technology**

Artificial intelligence used to evaluate job candidates must not become a tool that exacerbates discrimination.



October 2018

**Amazon scraps secret AI recruiting
tool that showed bias against women**



July 2015

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

Hiring before automation

Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination

September 2004

Marianne Bertrand

Sendhil Mullainathan

AMERICAN ECONOMIC REVIEW
VOL. 94, NO. 4, SEPTEMBER 2004
(pp. 991-1013)

We study race in the labor market by sending fictitious resumes to help-wanted ads in Boston and Chicago newspapers. To manipulate perceived race, resumes are randomly assigned African-American- or White-sounding names. **White names receive 50 percent more callbacks for interviews.** Callbacks are also more responsive to resume quality for White names than for African-American ones. The racial gap is uniform across occupation, industry, and employer size. We also find little evidence that employers are inferring social class from the names. Differential treatment by race still appears to still be prominent in the U. S. labor market.

discussion

Describe a use case

what are the **goals** of the AI system?

what are the **benefits** and to **whom**?

what are the **harms** and to **whom**?

Use case: Staples discounts

THE WALL STREET JOURNAL.

December 2012

WHAT THEY KNOW

Websites Vary Prices, Deals Based on Users' Information

By Jennifer Valentino-DeVries, Jeremy Singer-Vine and Ashkan Soltani

December 24, 2012

WHAT PRICE WOULD YOU SEE?



It was the same Swingline stapler, on the same Staples.com website. But for Kim Wamble, the price was \$15.79, while the price on Trude Frizzell's screen, just a few miles away, was \$14.29.

A key difference: where Staples seemed to think they were located.

A Wall Street Journal investigation found that the Staples Inc. website displays different prices to people after estimating their locations. More than that, **Staples appeared to consider the person's distance from a rival brick-and-mortar store**, either OfficeMax Inc. or Office Depot Inc. If rival stores were within 20 miles or so, Staples.com usually showed a discounted price.

<https://www.wsj.com/articles/SB10001424127887323777204578189391813881534>

Use case: AdFisher

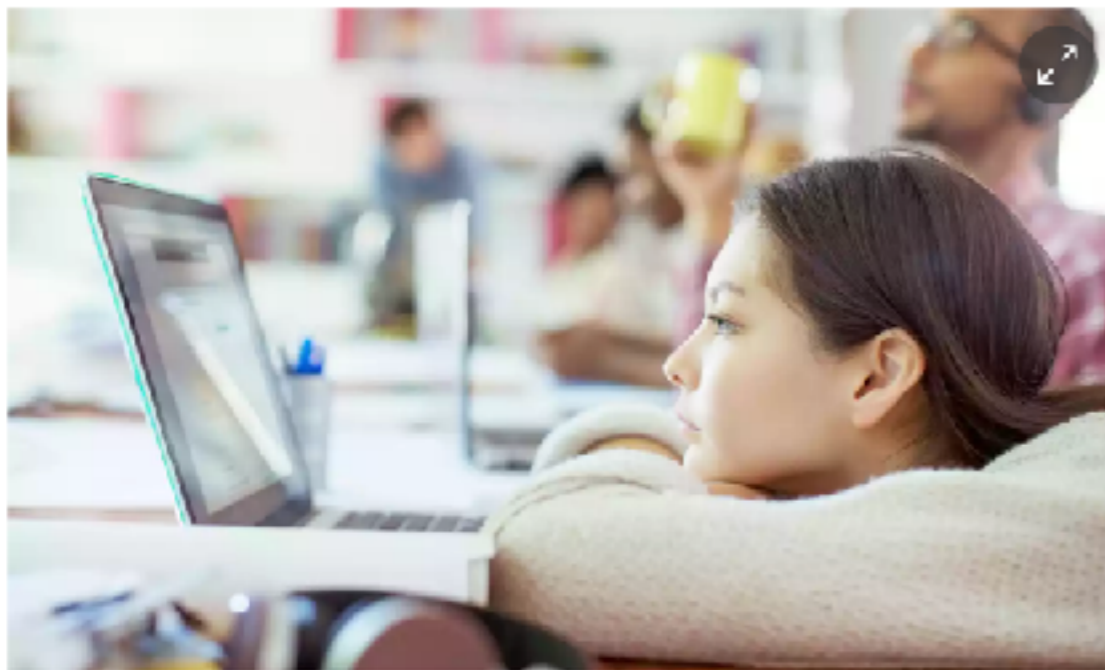
theguardian

July 2015

Samuel Gibbs

Wednesday 8 July 2015 11.29 BST

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



One experiment showed that Google displayed adverts for a career coaching service for executive jobs 1,852 times to the male group and only 318 times to the female group. Photograph: Alamy

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

The AdFisher tool simulated job seekers that did not differ in browsing behavior, preferences or demographic characteristics, except in gender.

One experiment showed that Google displayed ads for a career coaching service for “\$200k+” executive jobs **1,852 times to the male group and only 318 times to the female group.**

Another experiment, in July 2014, showed a similar trend but was not statistically significant.

<https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study>

Use case: Resume screening



Jeffrey Dastin

BUSINESS NEWS OCTOBER 9, 2018 / 11:12 PM / 6 MONTHS AGO

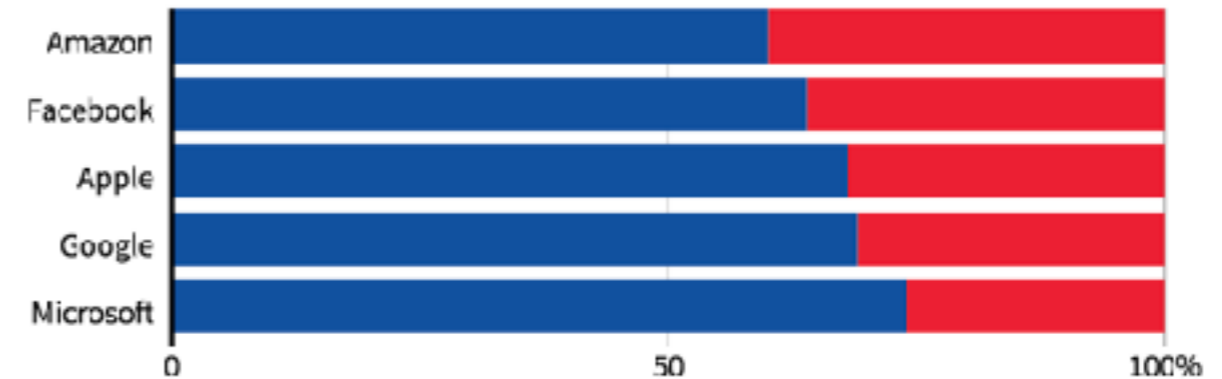
Amazon scraps secret AI recruiting tool that showed bias against women

“In effect, **Amazon’s system taught itself that male candidates were preferable**. It penalized resumes that included the word “women’s,” as in “women’s chess club captain.” And it **downgraded graduates of two all-women’s colleges**, according to people familiar with the matter. They did not specify the names of the schools.”

<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

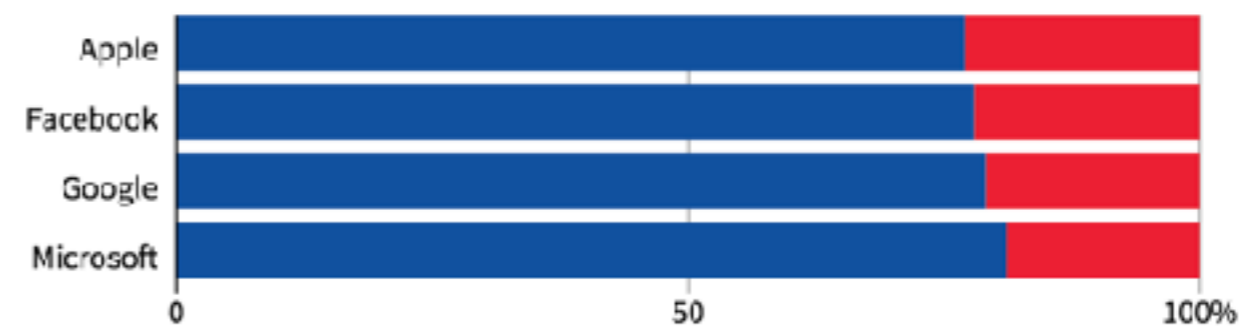
GLOBAL HEADCOUNT

Male Female



October 2018

EMPLOYEES IN TECHNICAL ROLES



“Note: Amazon does not disclose the gender breakdown of its technical workforce.”

Use case: Instant Checkmate

February 2013

Google
AdSense



Ads by Google

Latanya Sweeney, Arrested?

1) Enter Name and State. 2) Access F
Checks Instantly.

www.instantcheckmate.com/

Latanya Sweeney

Public Records Found For: Latanya S
www.publicrecords.com/

Latanya

A screenshot of the Instant Checkmate website. The header includes the 'instantcheckmate' logo and navigation links for 'DASHBOARD', 'EDIT ACCOUNT INFO', and 'LOGOUT'. The main profile section for 'LATANYA SWEENEY' includes a silhouette icon, address (1422 Centre Ave, Pittsburgh, PA 15219), and birth date (Oct 27, 1959, 53 years old). A 'CERTIFIED' badge is visible in the top right. Below the profile are sections for 'Personal' (Name, aliases, birthdate, phone numbers, etc.), 'Location' (Detailed address history and related data, maps, etc.), and 'Related Persons'. A 'Criminal History' section features a star rating and a disclaimer: 'This section contains possible citation, arrest, and criminal records for the subject of this report. While our database does contain hundreds of millions of arrest records, different counties have different rules regarding what information they will and will not release. We share with you as much information as we possibly can, but a clean slate here should not be interpreted as a guarantee that Latanya Sweeney has never been arrested; it simply means that we were not able to locate any matching arrest records in the data that is available to us.' A 'View Details' button is located at the bottom right of the profile section.

Racism is Poisoning Online Ad Delivery, Says Harvard Professor

Google searches involving black-sounding names are more likely to serve up ads suggestive of a criminal record than white-sounding names, says computer scientist

racially identifying names trigger ads suggestive of a criminal record

<https://www.technologyreview.com/s/510646/racism-is-poisoning-online-ad-delivery-says-harvard-professor/>

Use case: Amazon same-day delivery

Bloomberg

Amazon Doesn't Consider the Race of Its Customers. Should It?

“... In six major same-day delivery cities, however, **the service area excludes predominantly black ZIP codes** to varying degrees, according to a Bloomberg analysis that compared Amazon same-day delivery areas with U.S. Census Bureau data.”

<https://www.bloomberg.com/graphics/2016-amazon-same-day/>

New York City



Use case: Amazon same-day delivery

Bloomberg

Amazon Doesn't Consider the Race of Its Customers. Should It?

“The most striking gap in Amazon’s same-day service is in Boston, where **three ZIP codes encompassing the primarily black neighborhood of Roxbury are excluded** from same-day service, while the neighborhoods that surround it on all sides are eligible.”

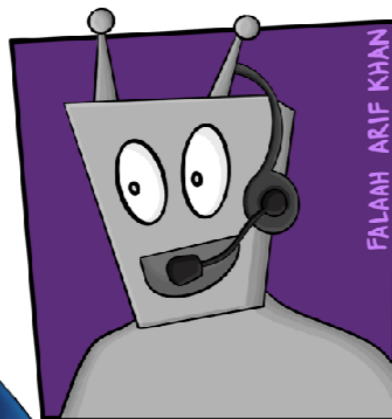
<https://www.bloomberg.com/graphics/2016-amazon-same-day/>



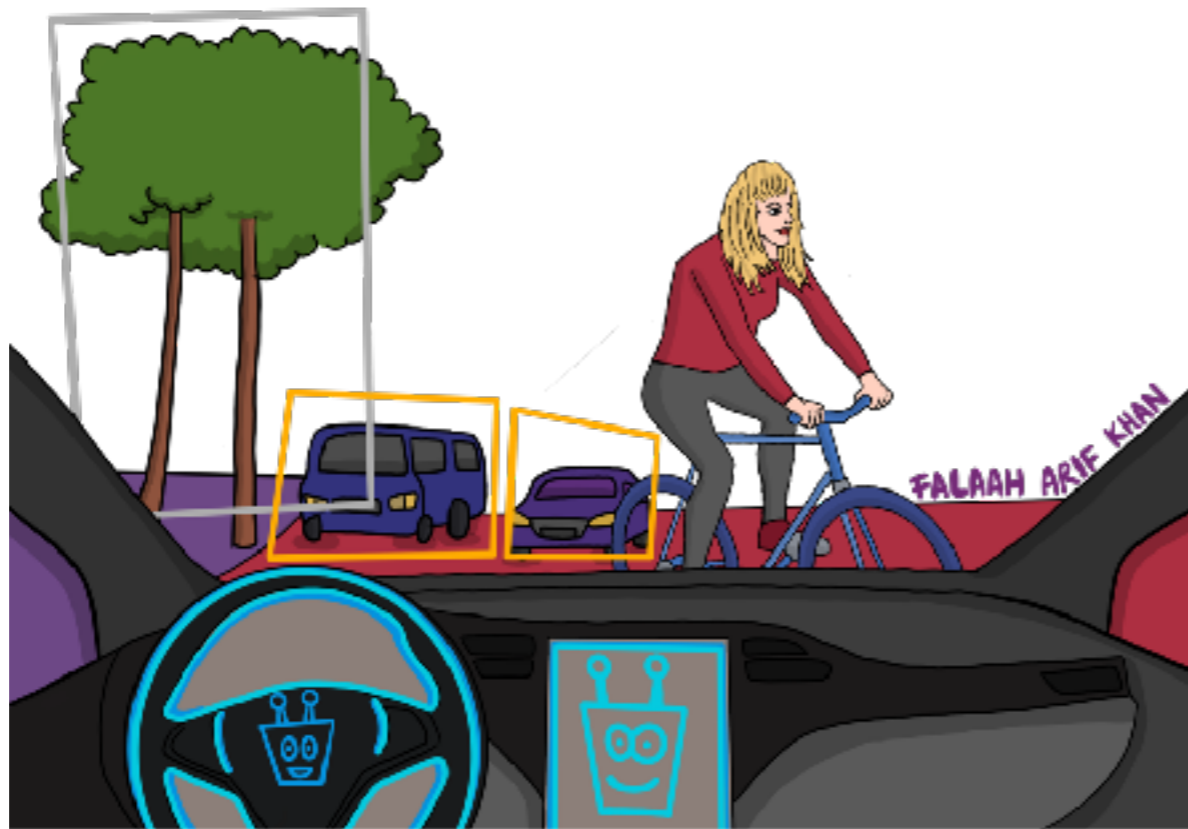


AI ethics teaser

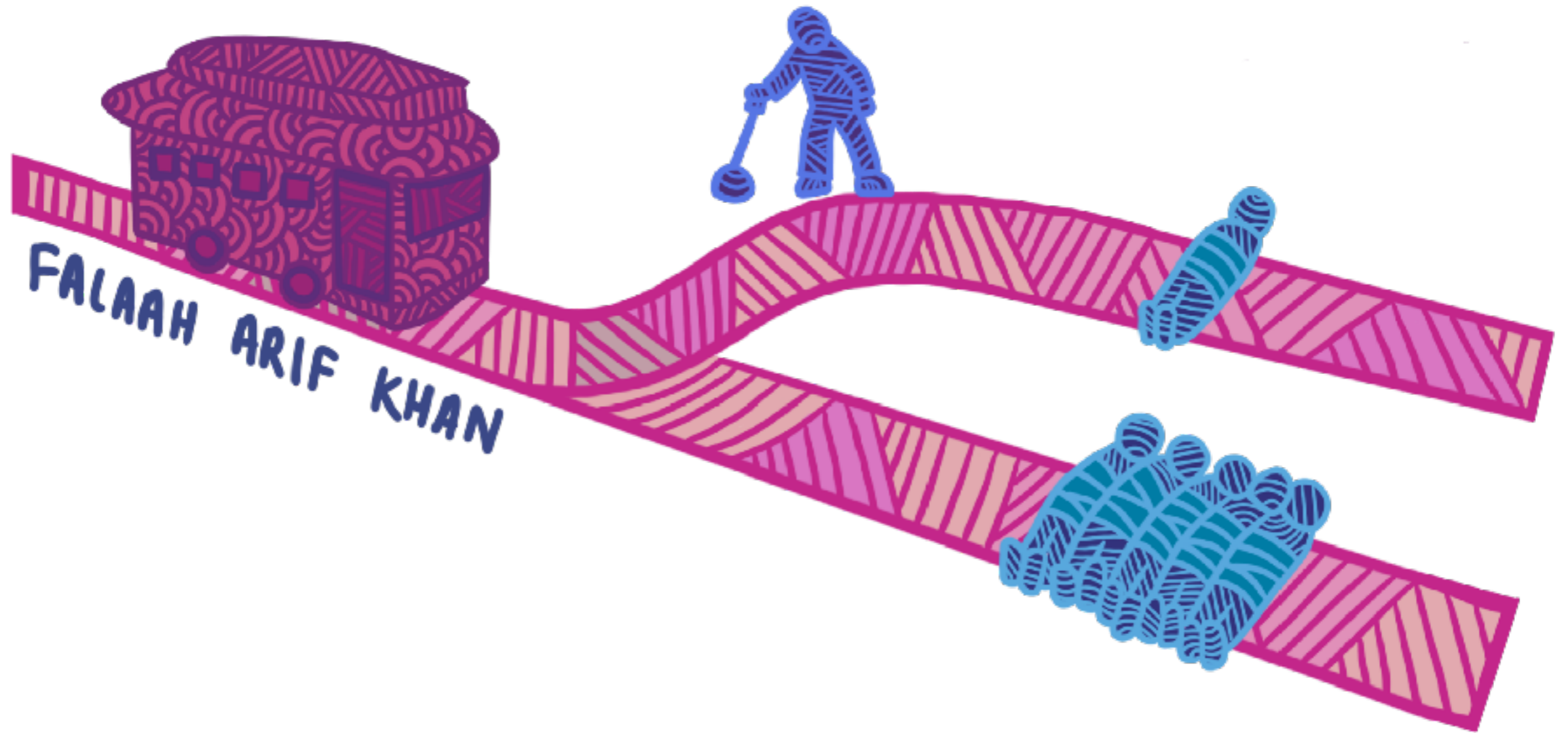
Mistakes lead to harms



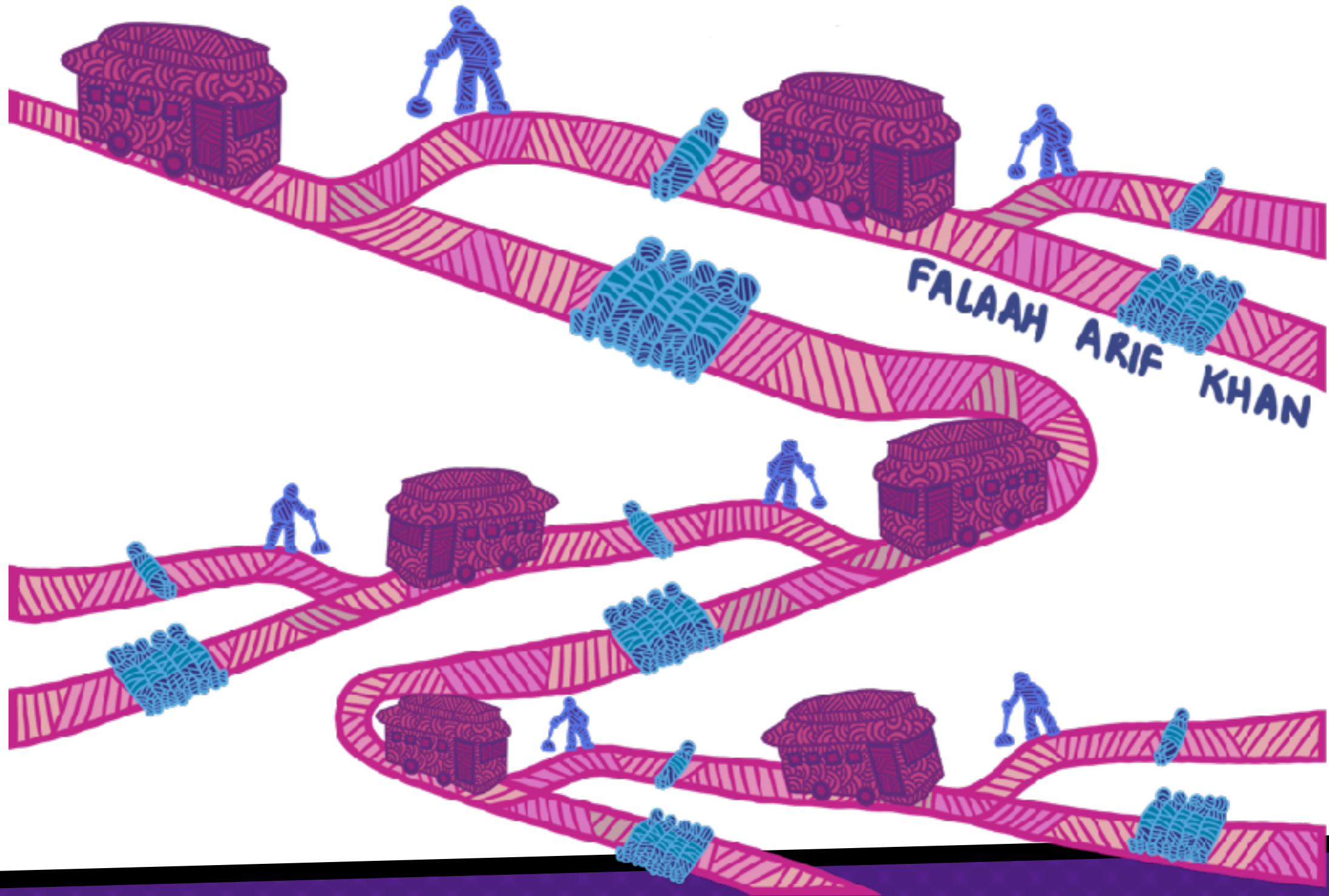
Mistakes lead to harms



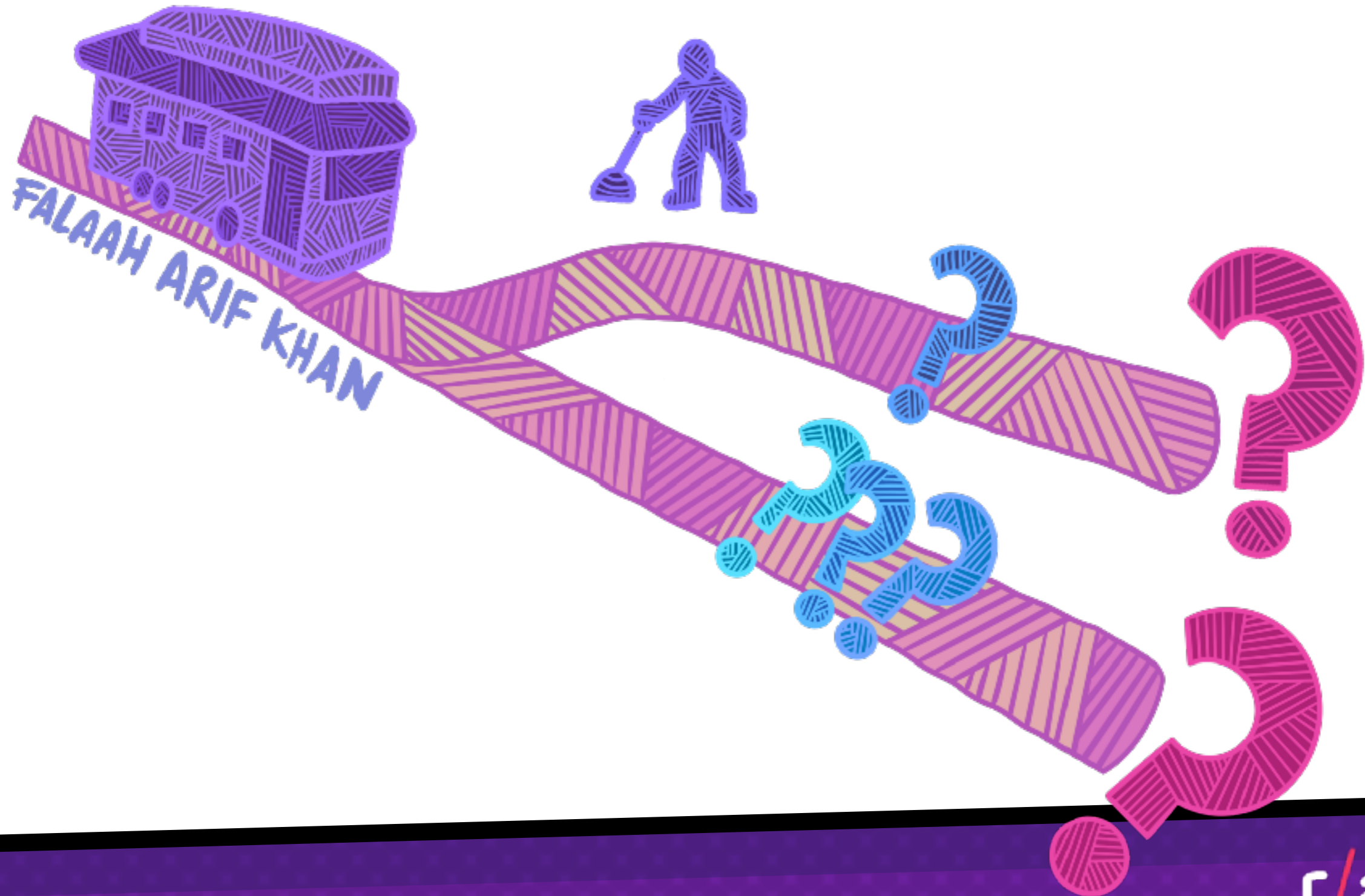
The trolley problem



The trolley problem



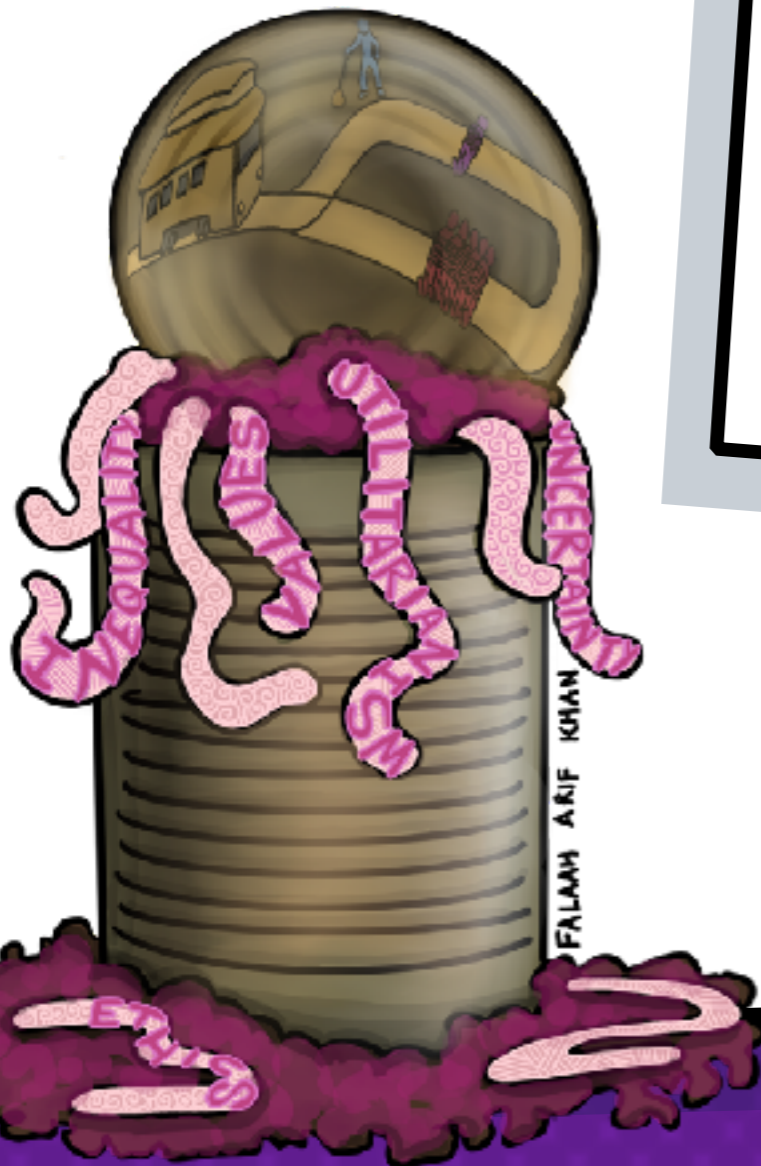
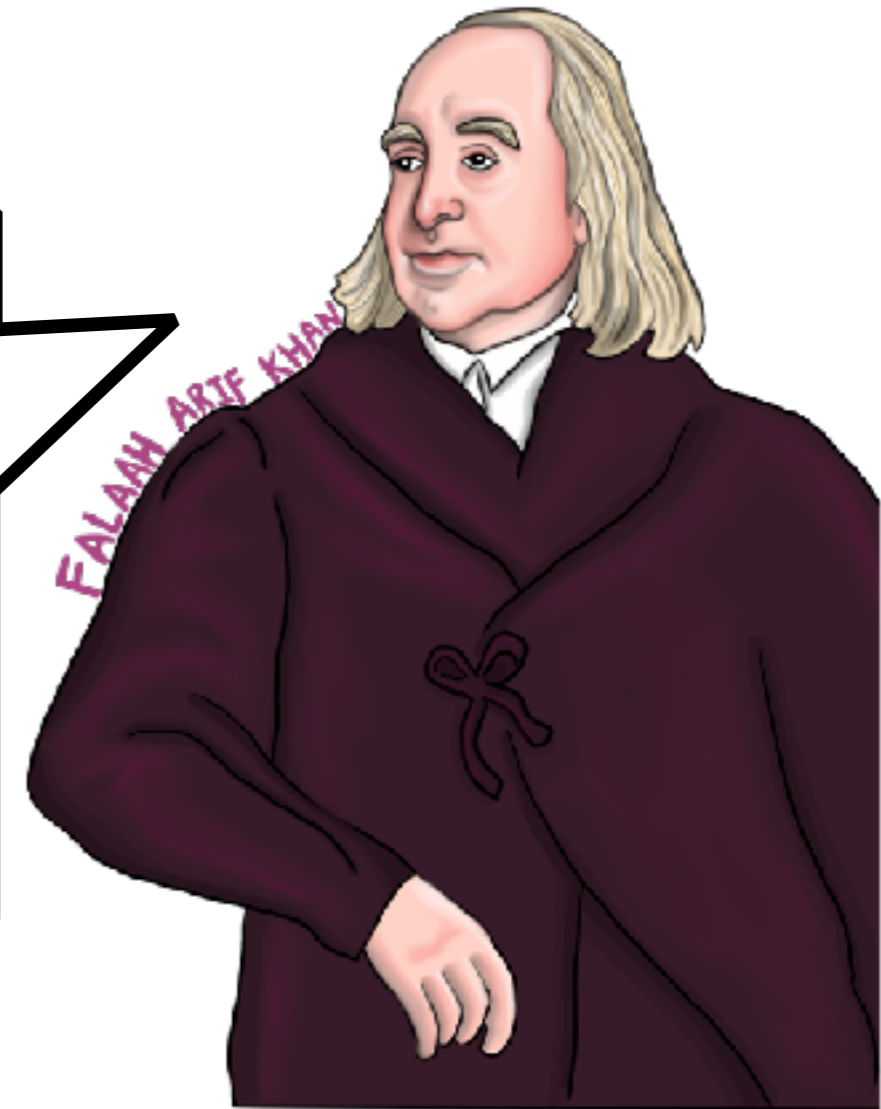
Dealing with uncertainty



Utilitarianism

“It is the greatest happiness of the greatest number that is the measure of right and wrong.”

Jeremy Bentham



Algorithmic morality?

Algorithmic morality

is the act of attributing moral reasoning to algorithmic systems



Algorithmic morality?



Tech rooted in people





**examples: racial
bias in risk
assessment**

Racial bias in criminal sentencing

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

May 23, 2016

May 2016

A commercial tool COMPAS automatically predicts some categories of future crime to assist in bail and sentencing decisions. It is used in courts in the US.

The tool correctly predicts recidivism **61% of the time.**

Blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend.

The tool makes **the opposite mistake among whites**: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes.



<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Racial bias in criminal sentencing

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
May 23, 2016

May 2016

A commercial tool COMPAS automatically predicts some categories of future crime to assist in bail and sentencing decisions. It is used in courts in the US.

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Racial bias in healthcare

Dissecting racial bias in an algorithm used to manage the health of populations

October 2019

Ziad Obermeyer^{1,2,*}, Brian Powers³, Christine Vogeli⁴, Sendhil Mullainathan^{5,*†}

† See all authors and affiliations

Science 25 Oct 2019:
Vol. 366, Issue 6464, pp. 447-453
DOI: 10.1126/science.aax2342

Science

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and **affecting millions of patients**, exhibits significant **racial bias**: **At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses**. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm **predicts health care costs rather than illness**, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, **despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise**. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

Fixing bias in algorithms?

The New York Times

By Sendhil Mullainathan

Dec. 6, 2019

ECONOMIC VIEW

Biased Algorithms Are Easier to Fix Than Biased People

Racial discrimination by algorithms or by people is harmful — but that's where the similarities end.



Tim Cook

December 2019

In one study published 15 years ago, **two people applied for a job**. Their résumés were about as similar as two résumés can be. One person was named Jamal, the other Brendan.

In a study published this year, **two patients sought medical care**. Both were grappling with diabetes and high blood pressure. One patient was black, the other was white.

Both studies documented **racial injustice**: In the first, the applicant with a black-sounding name got fewer job interviews. In the second, the black patient received worse care.

But they differed in one crucial respect. In the first, hiring managers made biased decisions. In the second, the culprit was a computer program.

<https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html>

Fixing bias in algorithms?

The New York Times

By Sendhil Mullainathan

December 2019

Dec. 6, 2019

ECONOMIC VIEW

Biased Algorithms Are Easier to Fix Than Biased People

Racial discrimination by algorithms or by people is harmful — but that's where the similarities end.



Tim Cook

Changing algorithms is easier than changing people: software on computers can be updated; the “wetware” in our brains has so far proven much less pliable.

[...] In a 2018 [paper](#) [...], I took a cautiously optimistic perspective and argued that **with proper regulation, algorithms can help to reduce discrimination.**

But the key phrase here is “proper regulation,” which we do not currently have.

We must ensure all the necessary inputs to the algorithm, including the data used to test and create it, are carefully stored. * [...] **We will need a well-funded regulatory agency with highly trained auditors to process this data.**

<https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html>



*a push for
regulation*

Automated Decision Systems (ADS)

Automated Decision Systems (ADS)

process data about people

help make consequential decisions

combine human & automated decision making

aim to improve **efficiency** and promote **equity**

are subject to **auditing** and **public disclosure**




Regulating ADS?

Precautionary



Nah! I'm fine!



The Anti-Elon 
@antiElon

Regulation rocks!

 2.3K  9.2K  126K

Risk-based



New York City Local Law 49 of 2018

January 11, 2018

An **Automated Decision System (ADS)** is a “computerized implementation of algorithms, including those derived from machine learning or other data processing or artificial intelligence techniques, which are used to make or assist in making decisions.”

Form task force that surveys the current use of ADS in City agencies and develops procedures for:

- requesting and receiving an **explanation** of an algorithmic decision affecting an individual (3(b))
- interrogating ADS for **bias and discrimination** against members of legally-protected groups (3(c) and 3(d))
- allowing the **public** to **assess** how ADS function and are used (3(e)), and archiving ADS together with the data they use (3(f))

ADS regulation in NYC: take 1



Principles

- using ADS **where** they promote innovation and efficiency in service delivery
- promoting **fairness, equity, accountability,** and **transparency** in the use of ADS
- reducing potential harm **across the lifespan** of ADS

New York City Local Law 144 of 2021



THE NEW YORK CITY COUNCIL

Corey Johnson, Speaker

December 11, 2021

This bill would require that a **bias audit** be conducted on an automated employment decision tool prior to the use of said tool. The bill would also require that candidates or employees that reside in the city **be notified about the use of such tools** in the assessment or evaluation for hire or promotion, as well as, **be notified about the job qualifications and characteristics that will be used** by the automated employment decision tool. Violations of the provisions of the bill would be subject to a civil penalty.



course overview



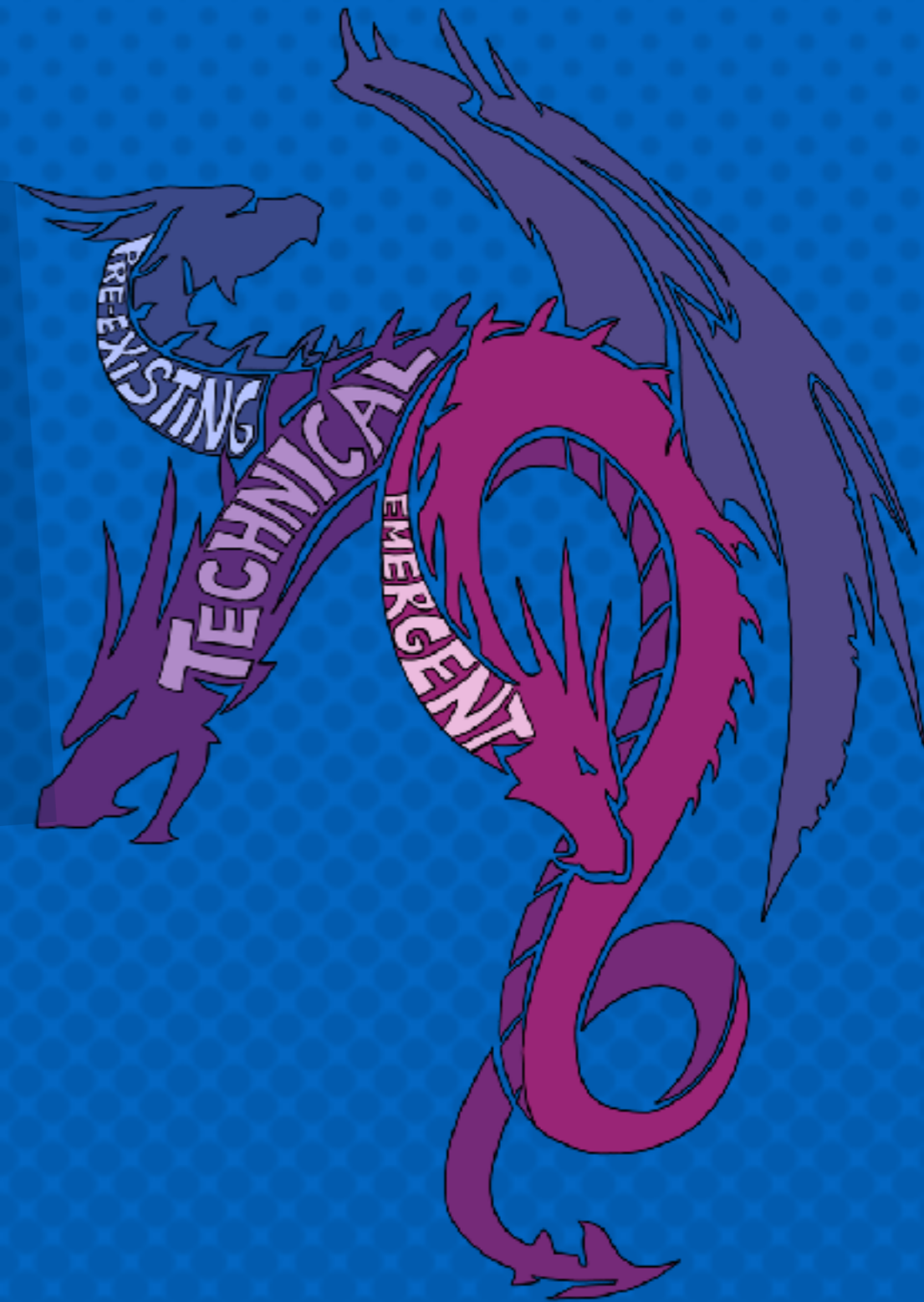
**module 1:
algorithmic
fairness**

Bias in computer systems

Pre-existing: exists independently of algorithm, has origins in society

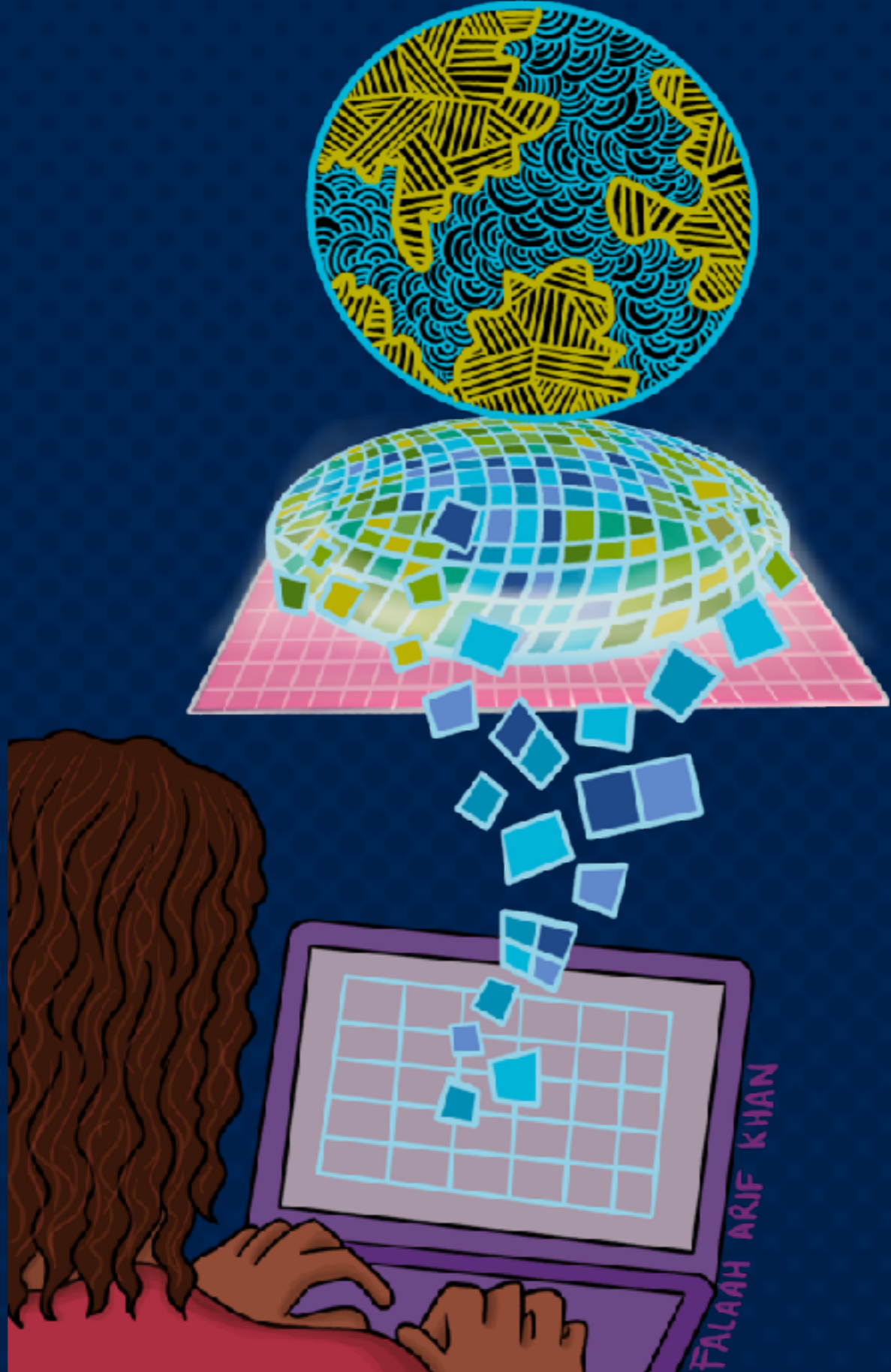
Technical: introduced or exacerbated by the technical properties of an ADS

Emergent: arises due to context of use



[Friedman & Nissenbaum (1996)]

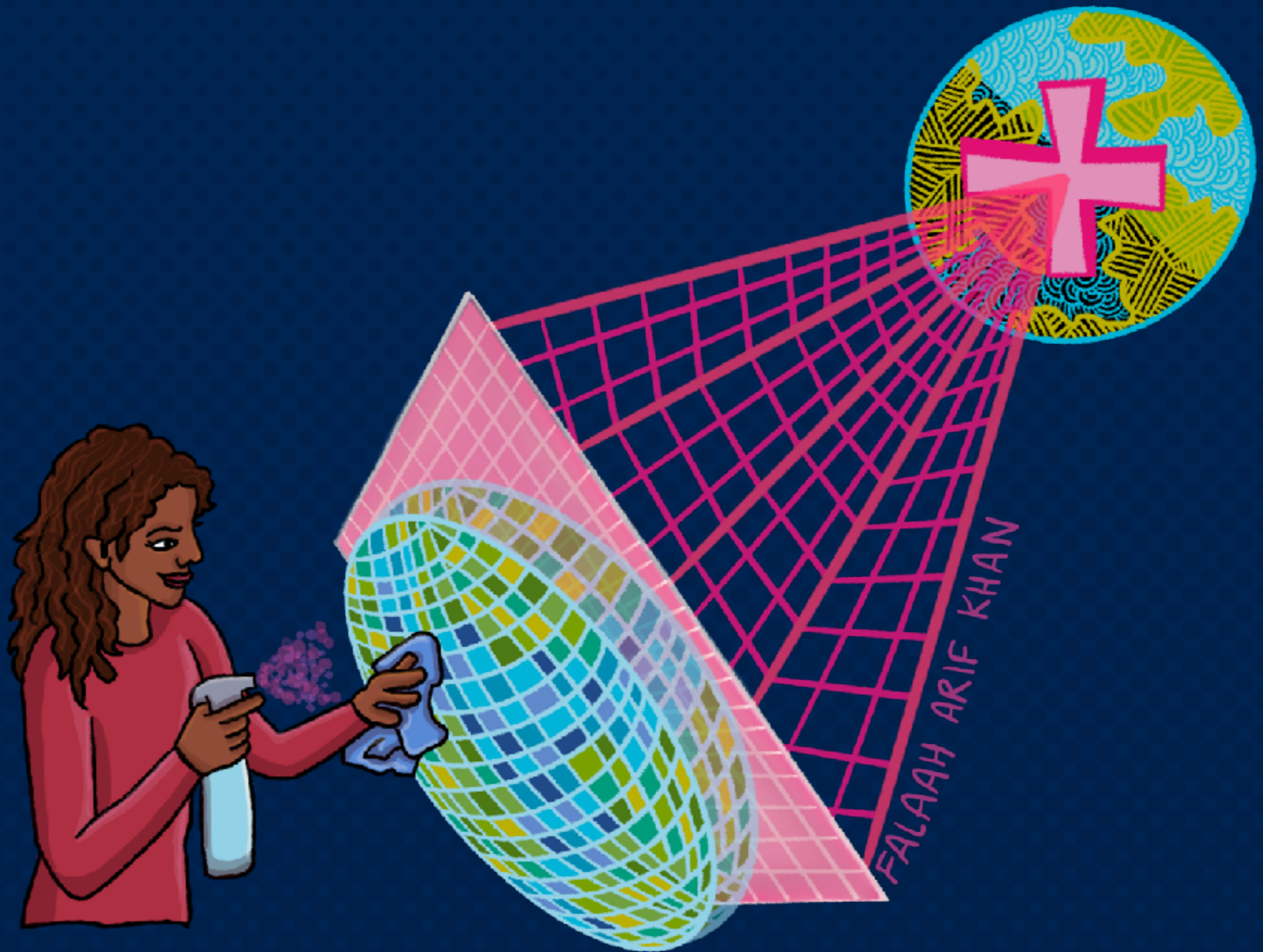




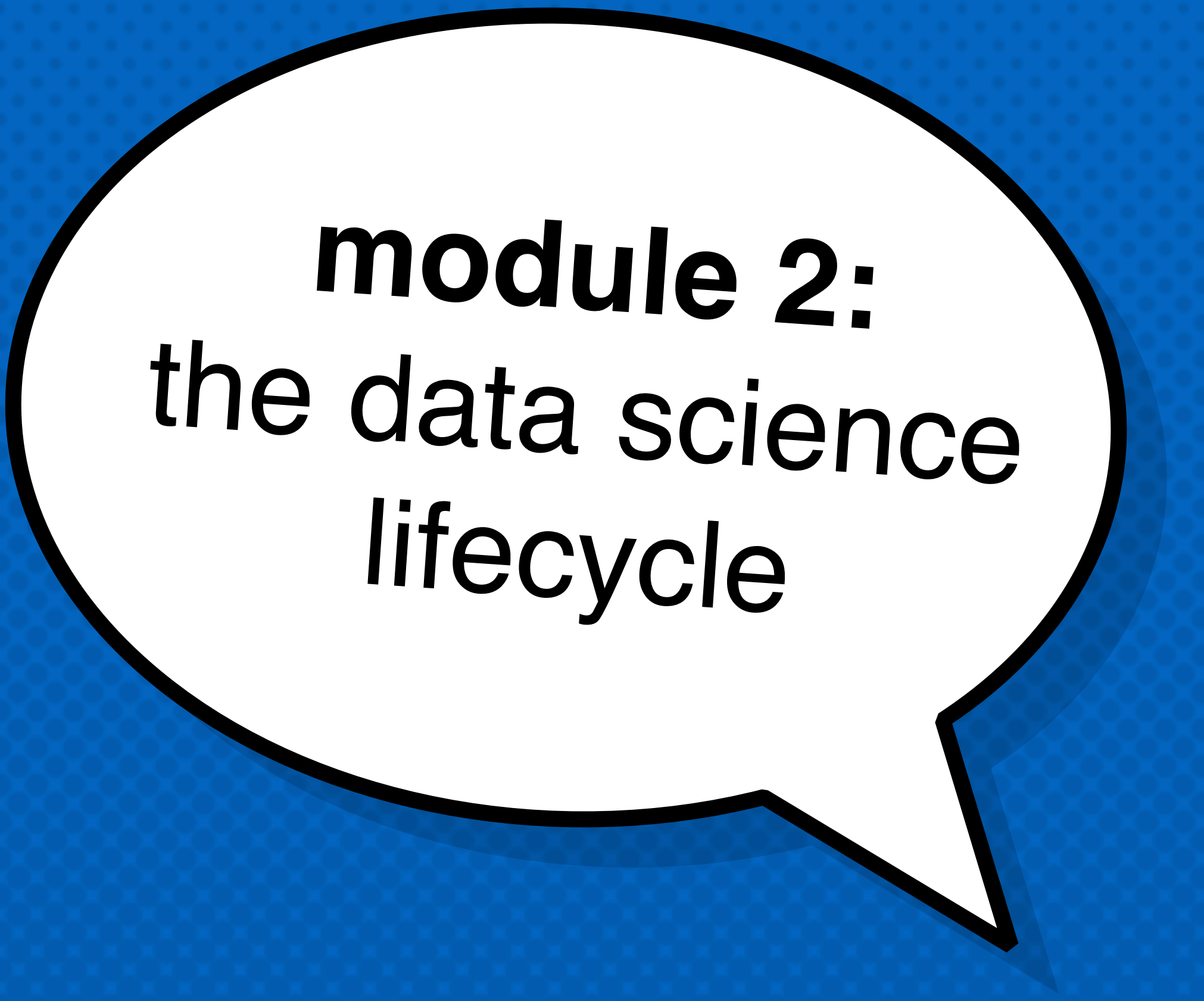
FALAH ARIF KHAN



FALAAH ARIF KHAN



FALAAH ARIF KHAN



module 2:
the data science
lifecycle

Bias in computer systems

Pre-existing: exists independently of algorithm, has origins in society

Technical: introduced or exacerbated by the technical properties of an ADS

Emergent: arises due to context of use

to fight bias, state
beliefs and
assumptions
explicitly

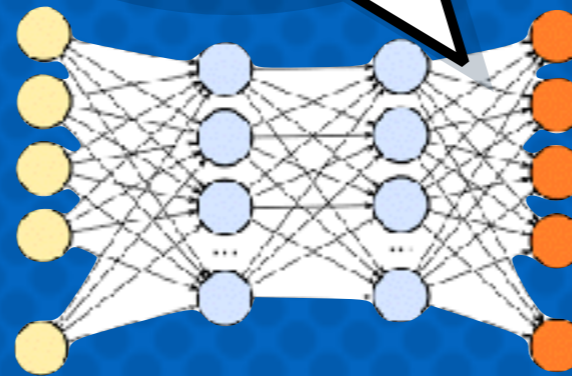


Fair-ML view

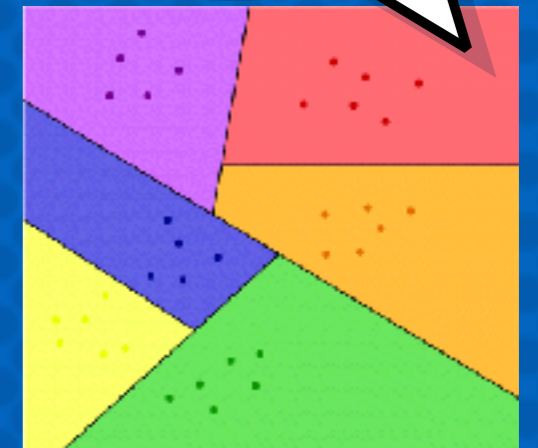
where did the data
come from?

id	sex	race	ethnicity	has_health_insurance	age	diabetes_mellitus	outcome_score
1	1	0	1	1	47.1/77	0	1
2	2	0	2	1	17.2/32	0	1
3	3	0	2	1	57.1/91	0	4
4	4	0	1	1	17.2/32	0	6
5	5	0	1	2	17.2/32	0	1
6	6	0	1	3	67.2/71	0	1
7	7	0	1	1	77.1/74	0	8
8	8	0	1	2	27.2/71	0	4
9	9	0	1	1	67.1/64	0	1
10	10	0	1	1	17.2/32	0	4
11	10	0	3	1	17.2/32	0	4
12	11	1	1	2	87.1/78	0	1
13	12	0	1	1	17.2/32	0	4
14	13	1	1	1	67.1/68	0	1
15	14	0	3	1	37.1/66	0	1
16	15	0	1	4	17.2/32	0	1
17	16	0	1	1	67.1/66	0	10
18	17	0	1	1	17.2/32	0	1
19	18	0	1	1	17.2/32	0	1
20	19	0	2	1	67.1/62	0	8
21	20	0	3	1	17.2/32	0	6
22	21	1	1	1	87.1/68	0	4
23	22	1	1	1	37.1/66	0	1
24	23	0	1	1	17.2/32	0	4
25	24	0	1	3	17.2/32	0	1
26	25	0	1	1	67.1/63	0	1
27	26	0	1	1	17.2/32	0	1
28	27	1	1	1	17.2/32	0	1
29	28	1	1	1	17.2/32	0	1

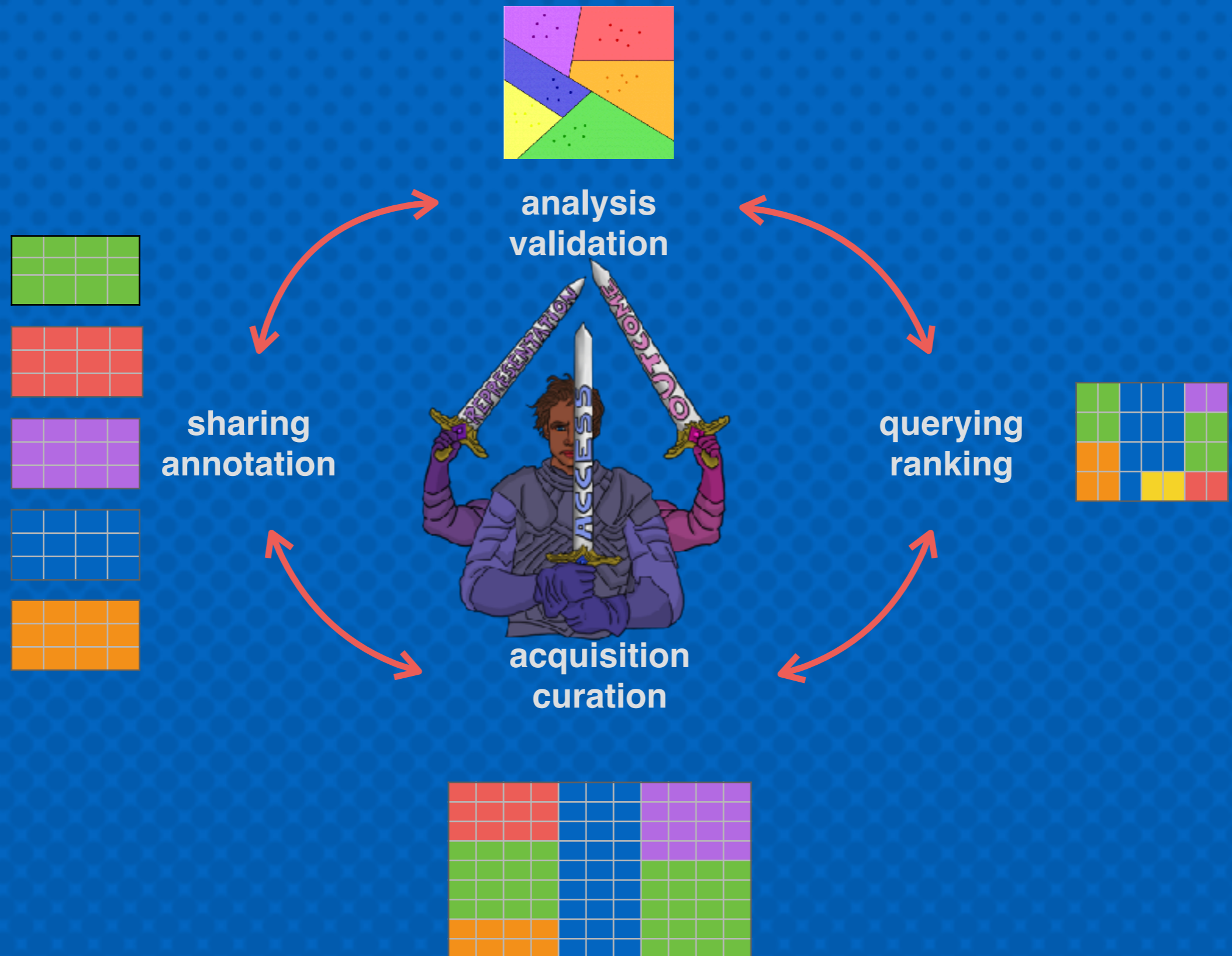
what happens
inside the box?



how are results
used?

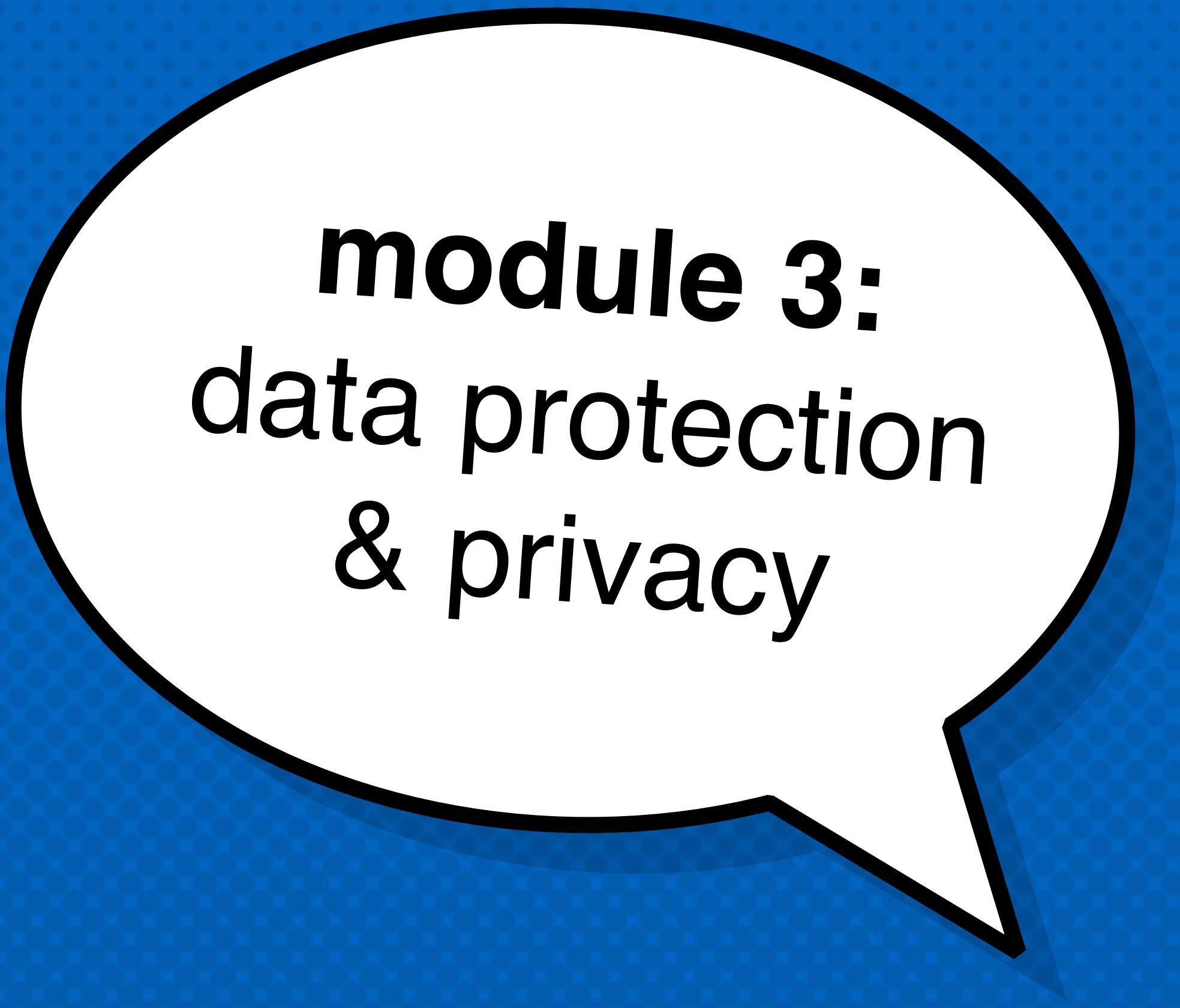


Lifecycle view



Models and assumptions





module 3:
data protection
& privacy

Privacy: two sides of the same coin

Did you go out drinking over the weekend?

protecting an individual

plausible deniability



learning about the population

noisy estimates

Truth or dare

Did you go out drinking over the weekend?

let's call this property **P** (Truth=Yes) and estimate **p**, the fraction of the group for whom **P** holds

thus, we estimate **p** as:

$$\tilde{p} = 2A - \frac{1}{2}$$

1. flip a coin **C1**
 1. if **C1** is tails, then **respond truthfully**
 2. if **C1** is heads, then flip another coin **C2**
 1. if **C2** is heads then **Yes**
 2. else **C2** is tails then respond **No**

} randomization - adding noise - is what gives plausible deniability a process privacy method

the expected number of **Yes** answers is:

$$A = \frac{3}{4}p + \frac{1}{4}(1-p) = \frac{1}{4} + \frac{p}{2}$$

← privacy comes from plausible deniability



Differential privacy

review articles

DOI:10.1145/1866739.1866758

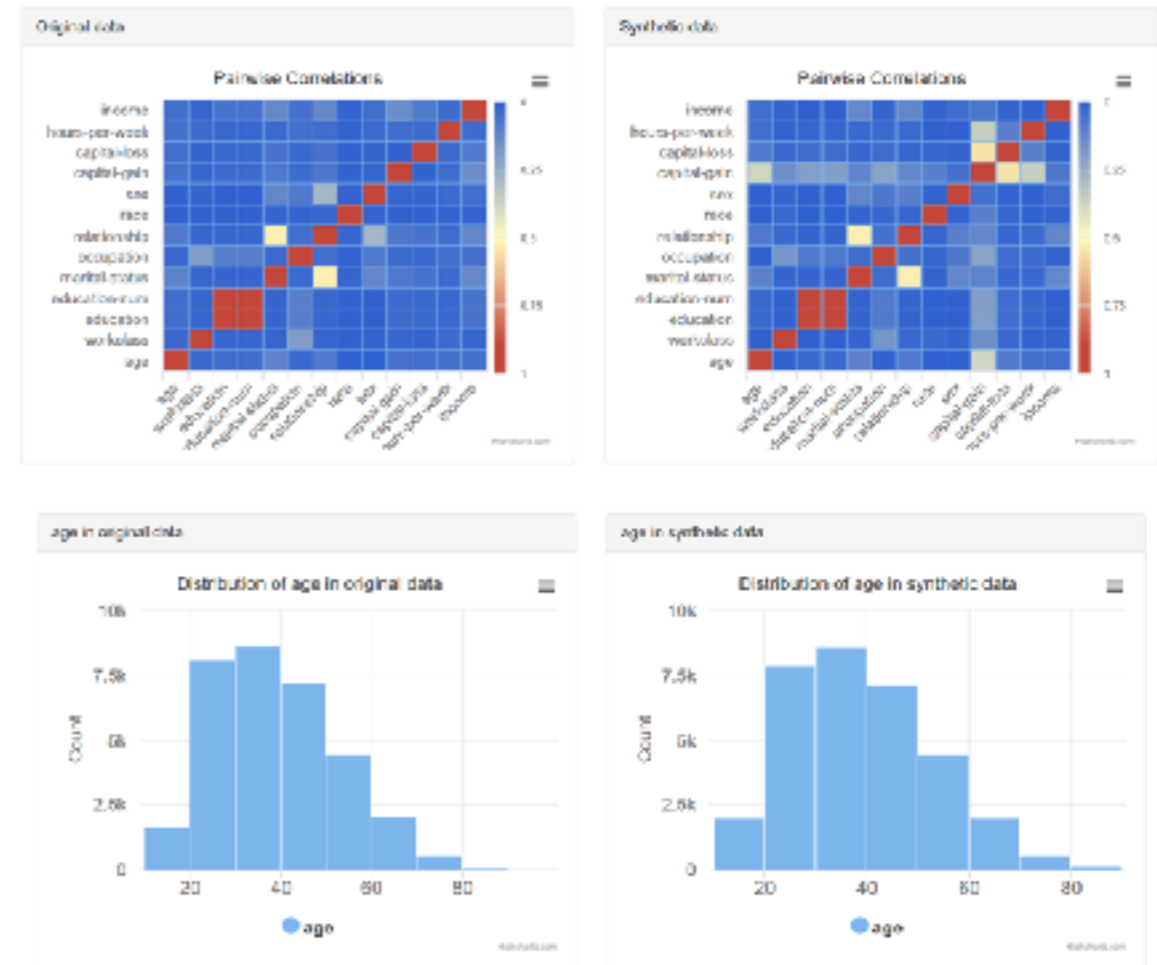
What does it mean to preserve privacy?

BY CYNTHIA DWORK

A Firm Foundation for Private Data Analysis

Communications of the ACM [CACM Homepage archive](#)

Volume 54 Issue 1, January 2011
Pages 86-95




Regulating ADS?

Precautionary



Nah! I'm fine!



The Anti-Elon 
@antiElon

Regulation rocks!

 2.3K  9.2K  126K

Risk-based



Legal frameworks

The screenshot shows the official website for the General Data Protection Regulation (GDPR). The navigation menu on the left lists various chapters and sections, including 'General provisions', 'Principles', 'Rights of the data subject', 'Controller and processor', 'Transfers of personal data to third countries or international organisations', 'Independent supervisory authorities', 'Cooperation and consistency', 'Remedies, liability and penalties', 'Provisions relating to specific processing situations', 'Delegated acts and implementing acts', and 'Final provisions'. The main content area features a 'Quick Access' section with links to each chapter.

Chapter	Articles
Chapter 1	1 2 3 4
Chapter 2	5 6 7 8 9 10 11
Chapter 3	12 13 14
Chapter 4	24 25 26
Chapter 5	44 45 46
Chapter 6	51 52 53
Chapter 7	60 61 62
Chapter 8	77 78 79
Chapter 9	85 86 87



Government
of Canada

Gouvernement
du Canada

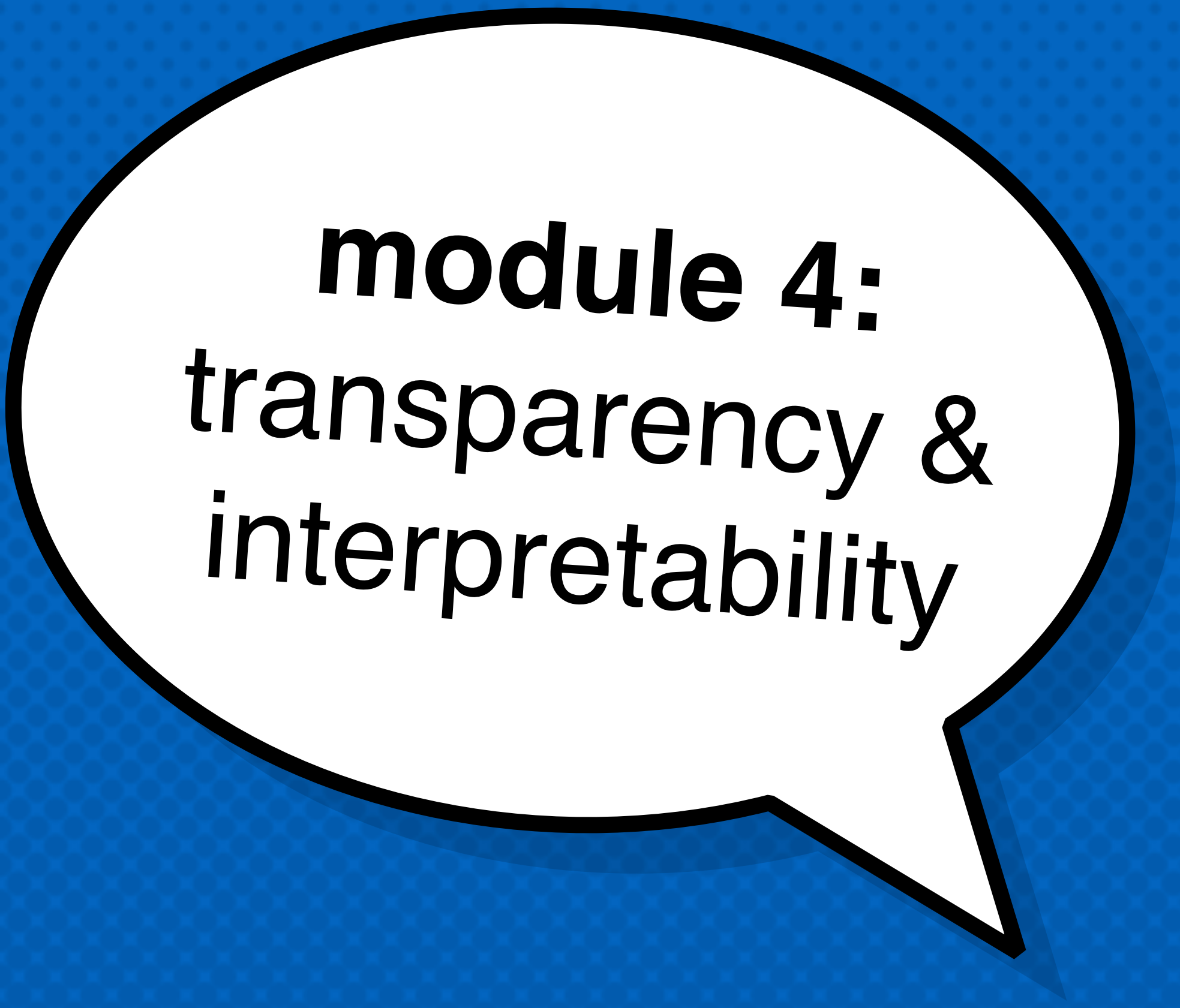


[Home](#) → [How government works](#) → [Policies, directives, standards and guidelines](#)

Directive on Automated Decision-Making

The Government of Canada is increasingly looking to utilize artificial intelligence to make, or assist in making, administrative decisions to improve service delivery. The Government is committed to doing so in a manner that is compatible with core administrative law principles such as transparency, accountability, legality, and procedural fairness. Understanding that this technology is changing rapidly, this Directive will continue to evolve to ensure that it remains relevant.

Date modified: 2019-02-05



module 4:
transparency &
interpretability

The evils of discrimination

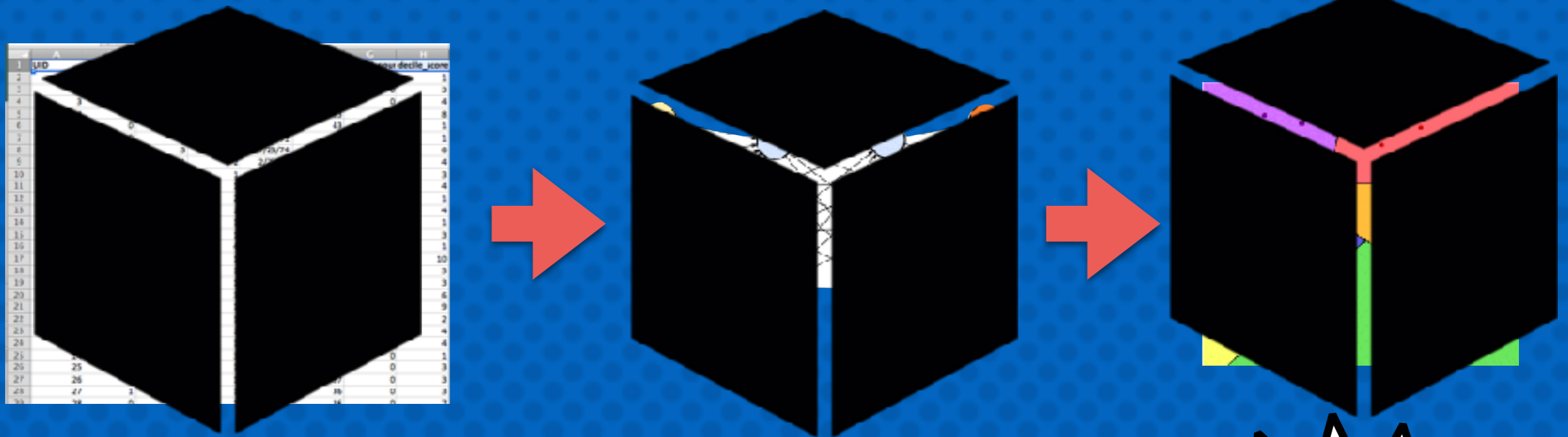
Disparate treatment

is the illegal practice of treating an entity, such as a job applicant or an employee, differently based on a **protected characteristic** such as race, gender, age, religion, sexual orientation, or national origin.

Disparate impact

is the result of systematic disparate treatment, where disproportionate **adverse impact** is observed on members of a **protected class**.

Regulating automated decisions

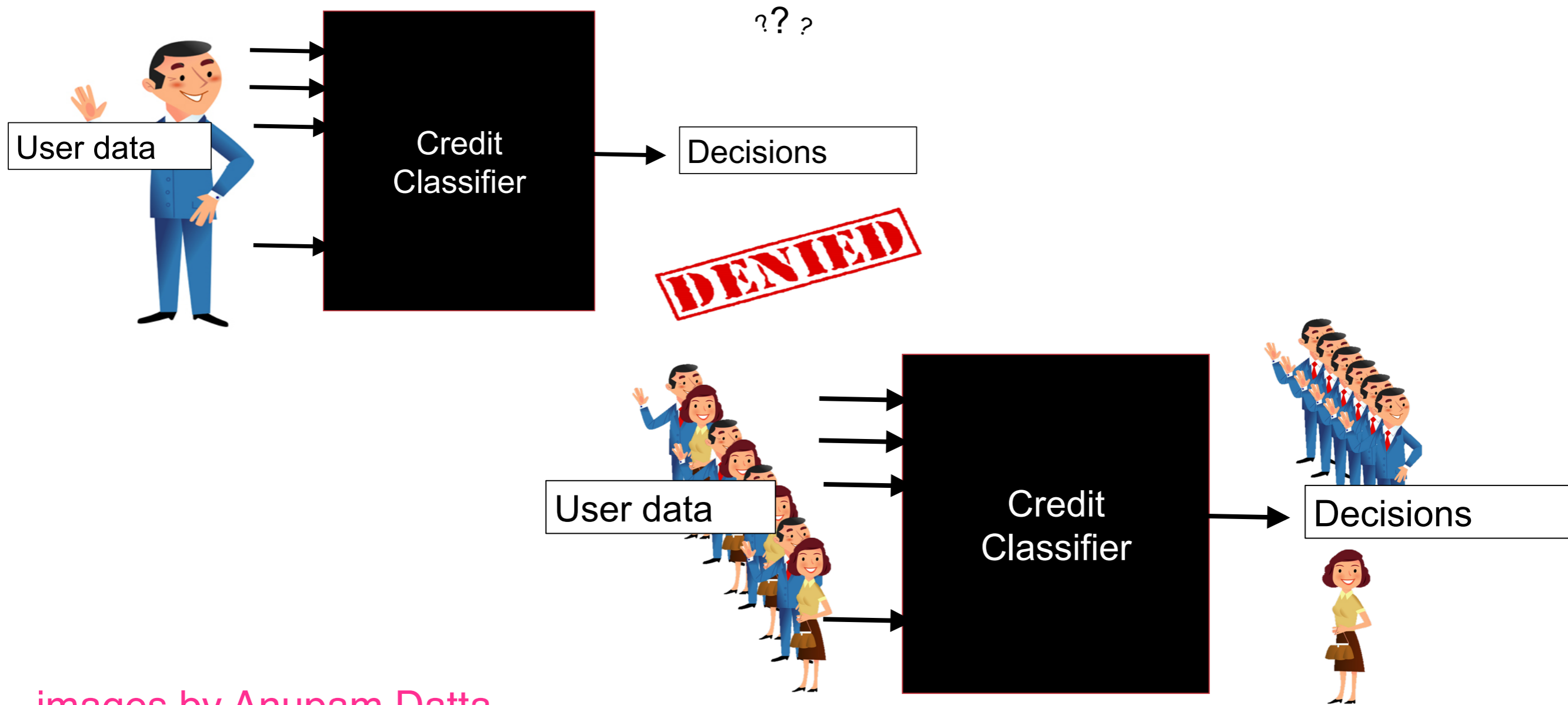


**Fair Housing
Act**

**Equal Credit
Opportunity
Act, 1964**

**Civil Rights
Act, 1964**

Auditing black-box models



images by Anupam Datta

Nutritional labels

Ranking Facts

Ingredients →

Attribute	Importance	
PubCount	1.0	
CSRankingAllArea	0.24	
Faculty	0.12	

Importance of an attribute in a ranking is quantified by the correlation coefficient between attribute values and items scores, computed by a linear regression model. Importance is high if the absolute value of the correlation coefficient is over 0.75, medium if this value falls between 0.25 and 0.75, and low otherwise.

Diversity overall ?

DeptSizeBin = Regional Code =

● Large ● Small ● NE ● W ● MW ● SA ● SC

Fairness ? →

DeptSizeBin	FA*IR	Pairwise	Proportion
Large	Fair	Fair	Fair
Small	Unfair	Unfair	Unfair

A ranking is considered unfair when the p-value of the corresponding statistical test falls below 0.05.

← Stability

Top-K	Stability
Top-10	Stable
Overall	Stable

comprehensible: short, simple, clear

consultative: provide actionable info

comparable: implying a standard



in summary

So what is RDS?

As advertised: ethics, legal compliance, personal responsibility.
But also: **data quality!**

A technical course, with content drawn from:

1. fairness, accountability and transparency
2. data engineering
3. security and privacy



We will learn **algorithmic techniques** for data analysis.
We will also learn about recent **laws / regulatory frameworks**.

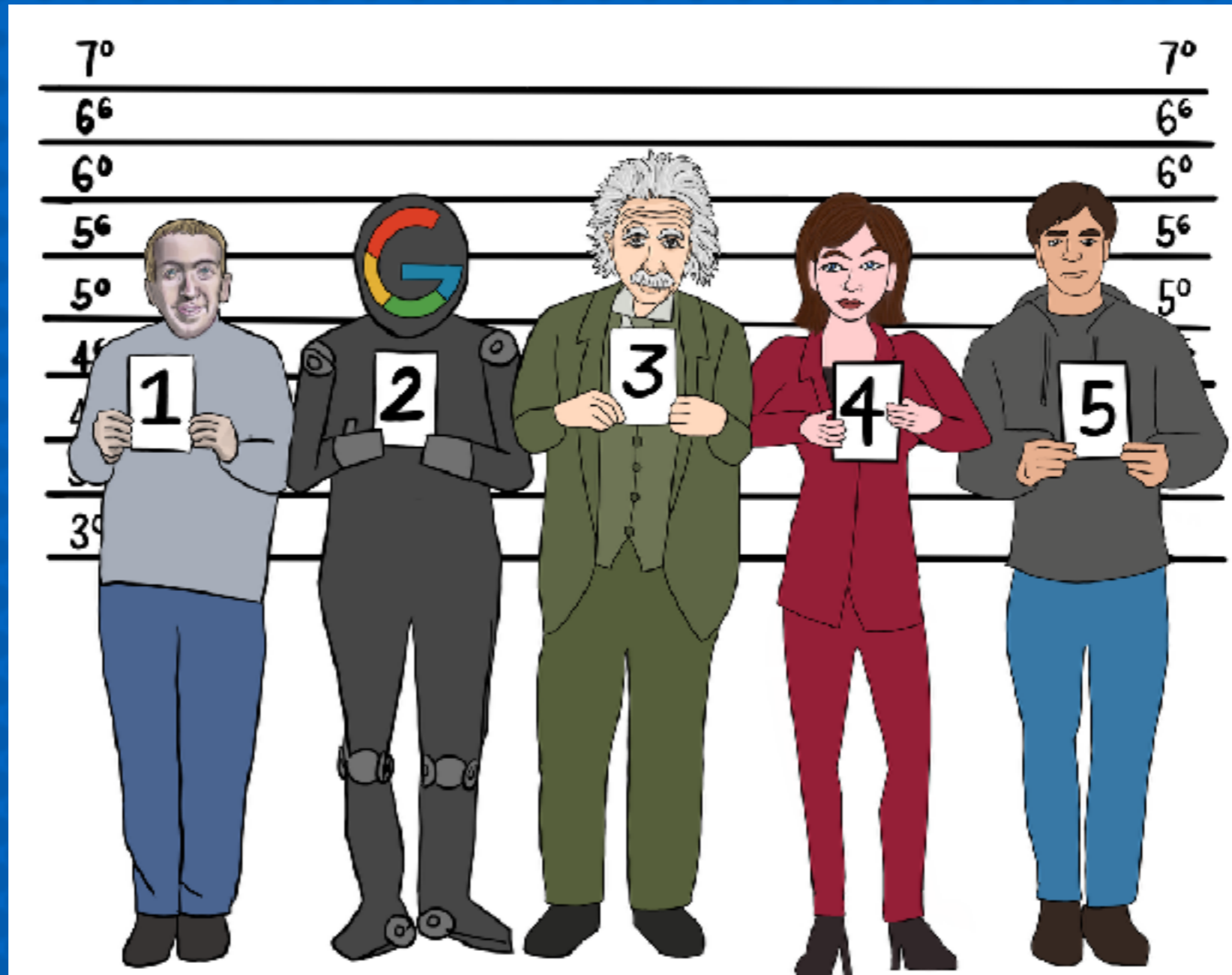
Bottom line: we will learn that many of the problems are **socio-technical**, and so cannot be “solved” with technology alone.

My perspective: a pragmatic engineer, **not** a technology skeptic.

Nuance, please!



We all are responsible



@FalaahArifKhan

Responsible Data Science

Introduction and Overview

Thank you!



NYU

TANDON SCHOOL
OF ENGINEERING



NYU

Center for
Data Science

r/ai