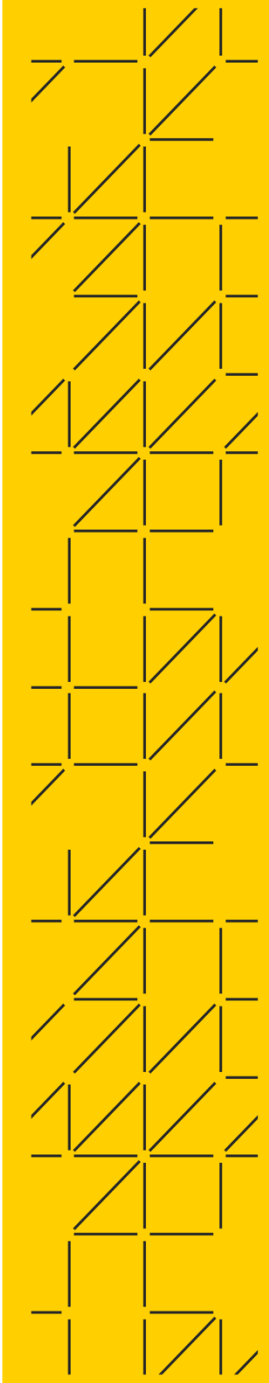


Edinburgh Napier
UNIVERSITY

School of Computing,
Engineering & the
Built Environment

Computing in Contemporary Society

Last week

- We examined the interaction between technology and democracy, focusing on access to information

This week

- We explore algorithms and algorithmic bias

Edinburgh Napier UNIVERSITY
School of Computing, Engineering & the Built Environment

Algorithms and Algorithmic Bias

Lecture five

Khristin Fabian

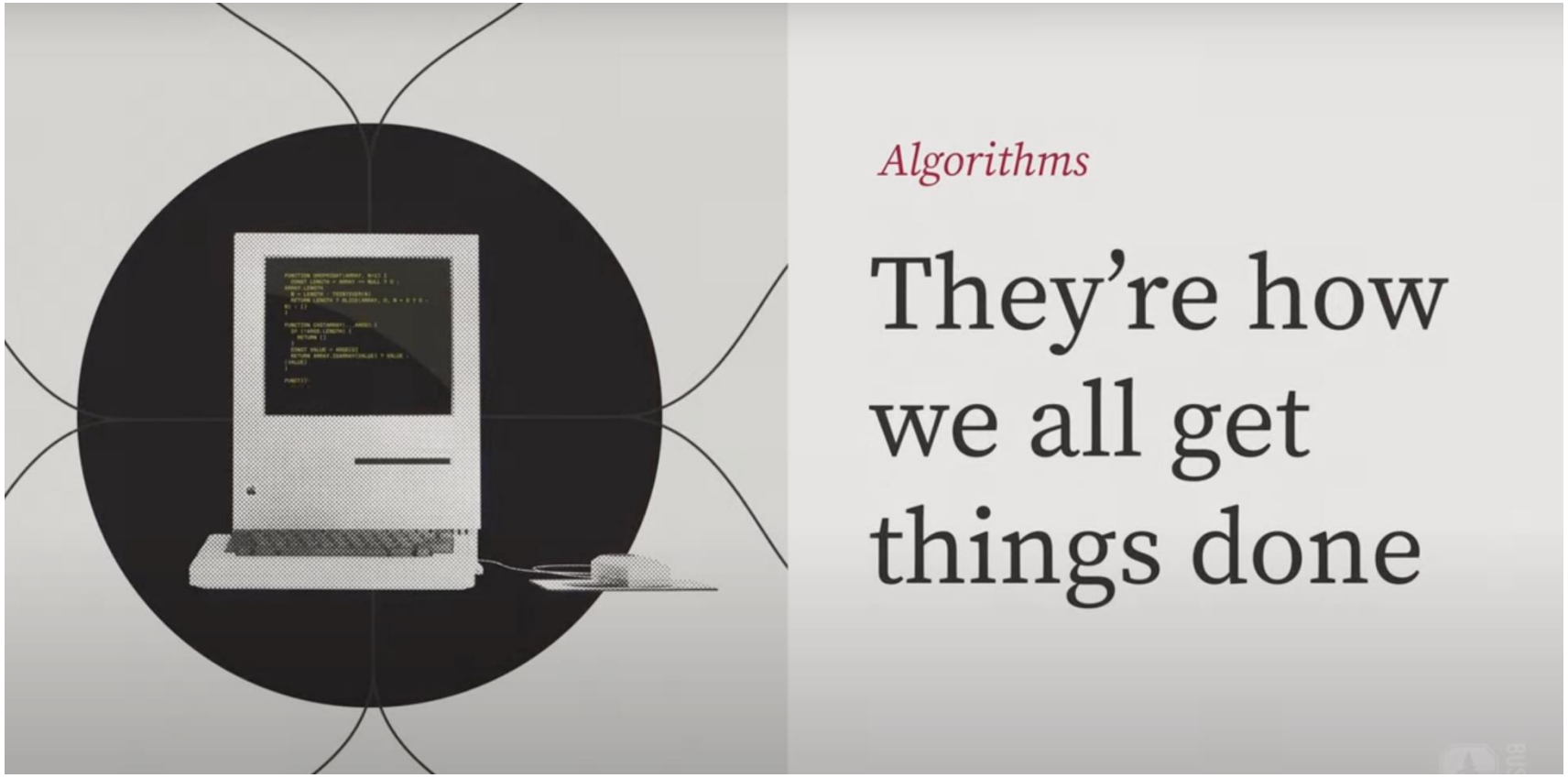
This week's reading

- Try the activity, "How Normal Am I" https://www.hownormalami.eu
- Read: How have social media algorithms changed the way we interact?, October 2024
- Read: AI hiring tools may be filtering out the best applicants, Feb 2024
- Read: How to avoid algorithmic decision-making mistakes: lessons from the Robodebt debacle https://stories.uq.edu.au/momentum-magazine/robodebt-algorithmic-decision-making-mistakes/index.html

Today's lecture is about

- Algorithms
- Algorithmic Bias

What is an algorithm



Activity: Algorithm Literacy

- Dogruel, Masur and Joeckel, 2021
- The instrument measures
  - awareness of algorithms use
  - knowledge about algorithms

https://forms.office.com/e/S2GXgnxN2q

AI Algorithms and ML Algorithms

- AI algorithms provide instructions for AI technology to think and react to data in ways that are intuitive to how we process information
- A machine learning algorithm is a set of rules or processes used by an AI system to conduct tasks—most often to discover new data insights and patterns, or to predict output values from a given set of input variables. Algorithms enable machine learning (ML) to learn. (IBM)

Join at menti.com | use code 2849 7607

What are some ways that ML or AI algorithms influence your life, whether large or small?

0 responses with votes

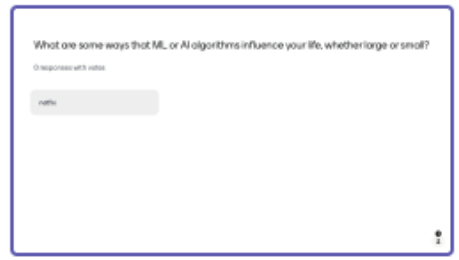
netflix

1

Menti

CCS Week 5

Choose a slide to present



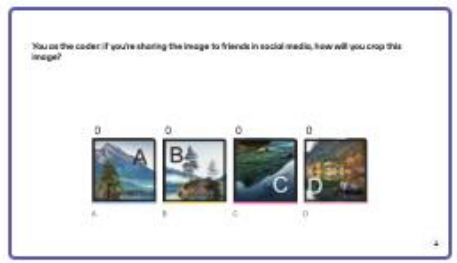
Join at menti.com | use code 2849 7607

You as the coder: if you're sharing the image to friends in social media, how will you crop this image?

Menti

CCS Week 5

Choose a slide to present



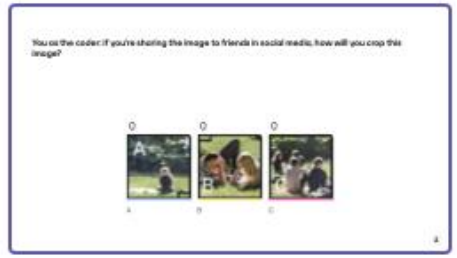
Join at menti.com | use code 2849 7607

You as the coder: if you're sharing the image to friends in social media, how will you crop this image?

Menti

CCS Week 5

Choose a slide to present



Join at menti.com | use code 2849 7607

You as the coder: if you're sharing the image to friends in social media, how will you crop this image?

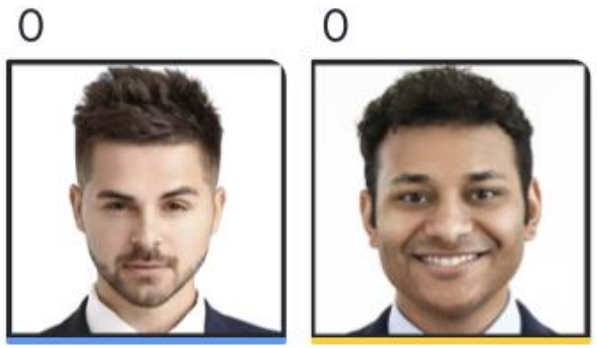
Menti

CCS Week 5

Choose a slide to present

Join at menti.com | use code 2849 7607

You as the coder: if you're sharing the image to friends in social media, how will you crop this image?

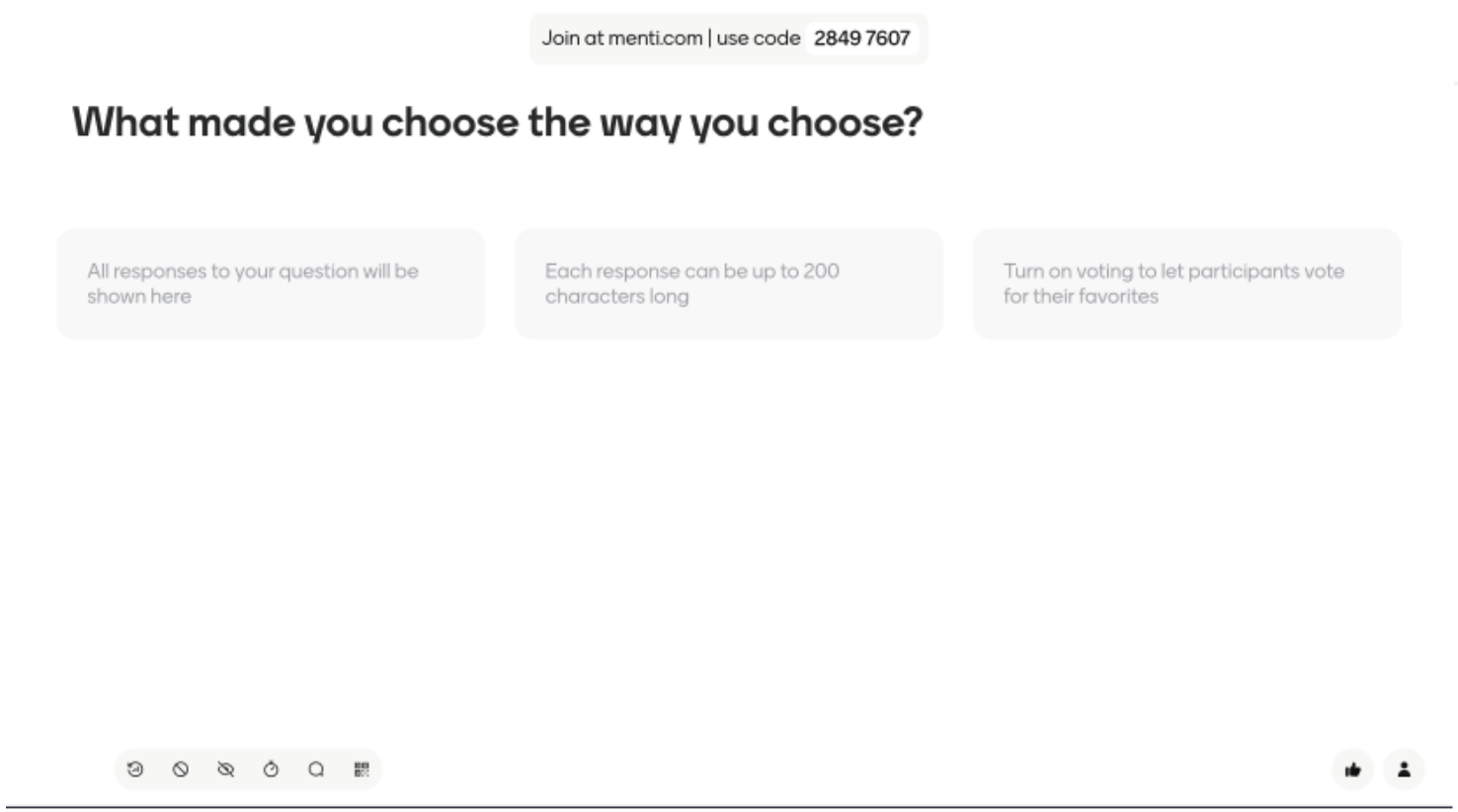


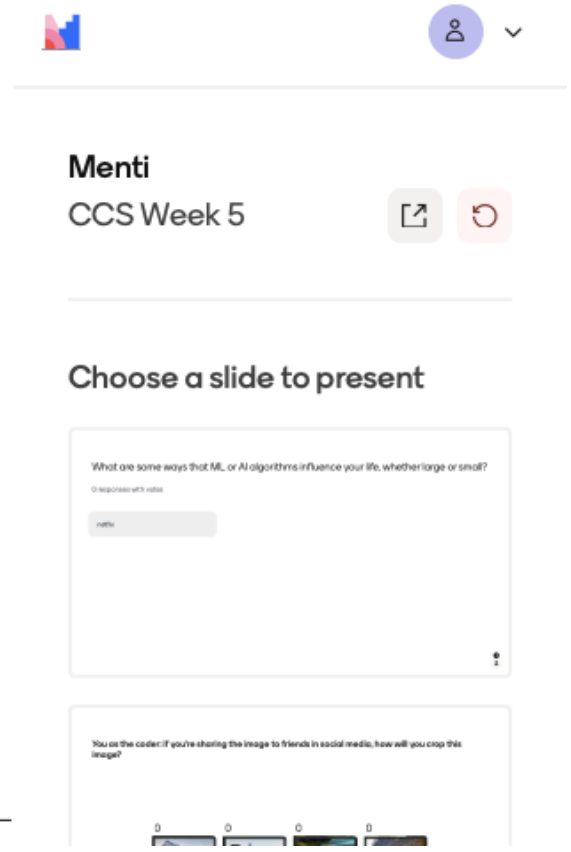
Menti

CCS Week 5

Choose a slide to present

What made you choose the way you





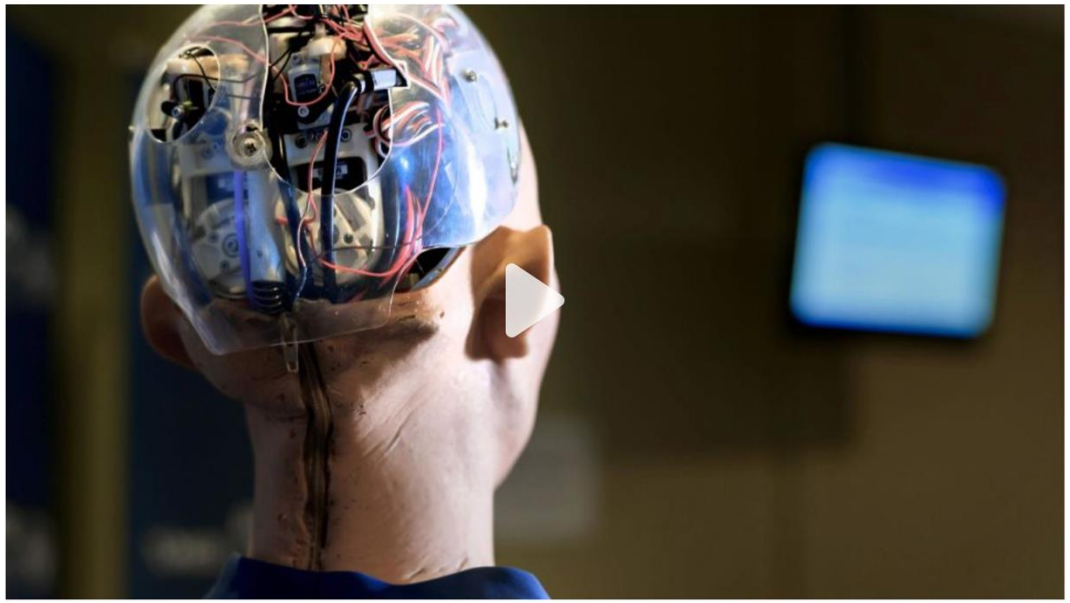
When algorithms choose…

- https://youtu.be/Ok5sKLXqynQ?feature=shared&t=84

Twitter's action…

Twitter says its image-cropping algorithm was biased, so it's ditching it

By Rachel Metz, CNN Business
4 minute read · Updated 7:49 PM EDT, Wed May 19, 2021



According to Twitter: The move away from using an algorithm to crop images lowers the company's dependency on machine learning for a function that we agree is best performed by people using our products.

https://edition.cnn.com/2021/05/19/tech/twitter-image-cropping-algorithm-bias/index.html

You as the "codee"

- Try the activity, ”How Normal Am I” https://www.hownormalami.eu [10 minutes]

You as the "codee"

- Try the activity, ”How Normal Am I” https://www.hownormalami.eu [8 minutes]
- The algorithms used in the program (except for the BMI) were ready made algorithms (algorithms developed by other people and shared on github). The choice of the algorithm was depending on its size (for loading purposes). What implication does this have?
- The face to body mass index algorithm doesn't make distinction between men and women, culture or fitness level. This example shows how algorithms can overlook critical nuances leading to potentially flawed conclusions.

Break (30 minutes)

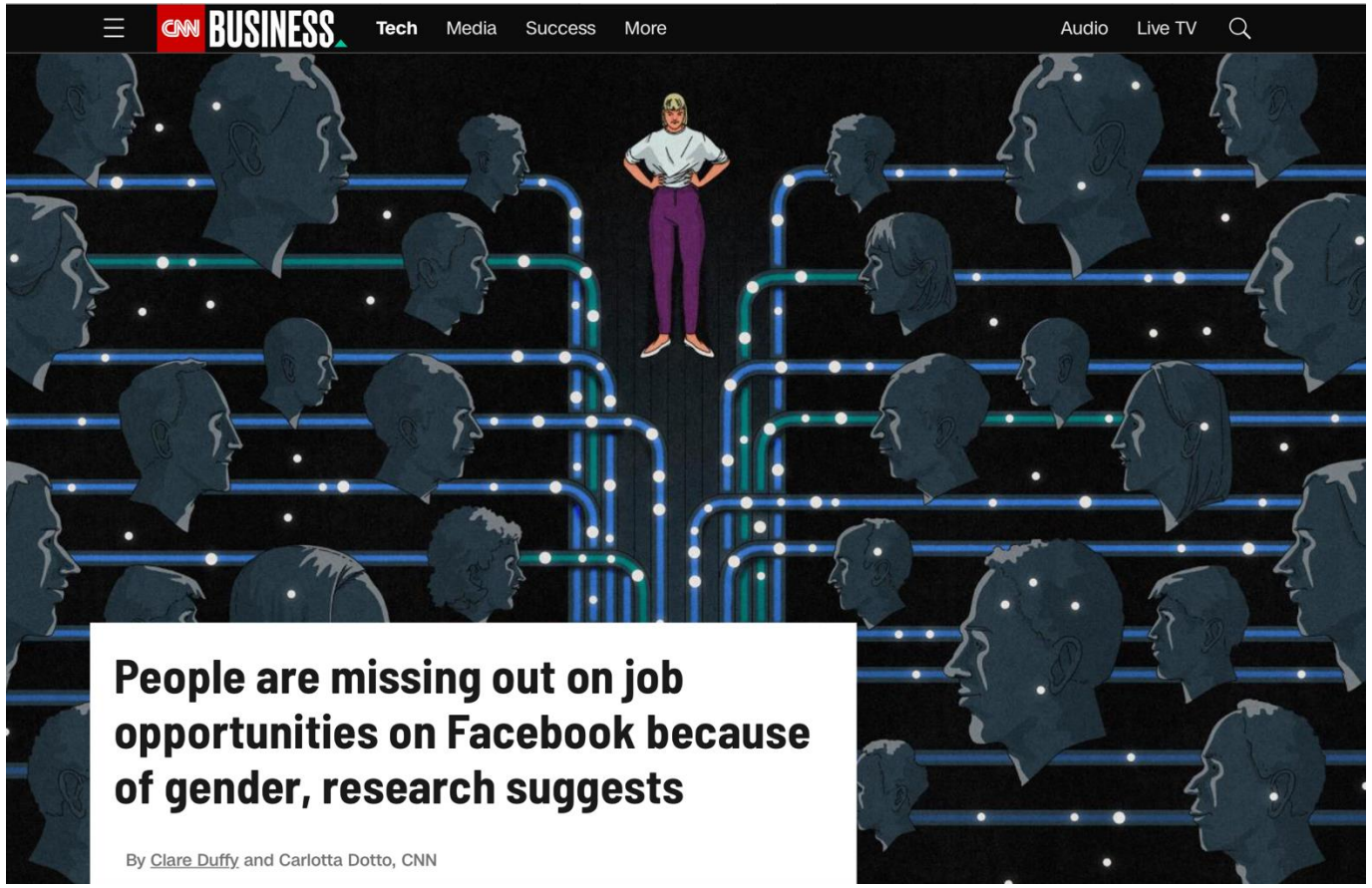
If you liked the website hownormalami, Tijmen Schep has a similar website called

Are you you? https://www.areyouyou.eu

In this game the goal is to beat the face recognition algorithm by making funny faces.

Re-convene at 11:30

Case study: Facebook ads algorithm and gender bias

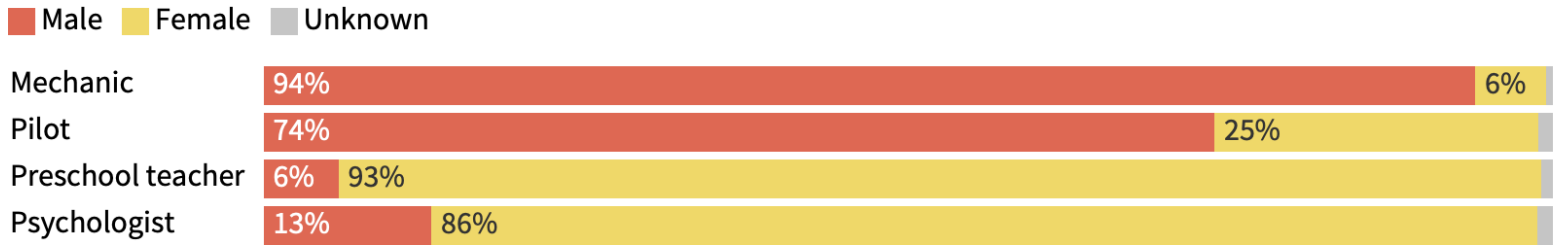


https://edition.cnn.com/2023/06/12/tech/facebook-job-ads-gender-discrimination-asequals-intl-cmd/index.html

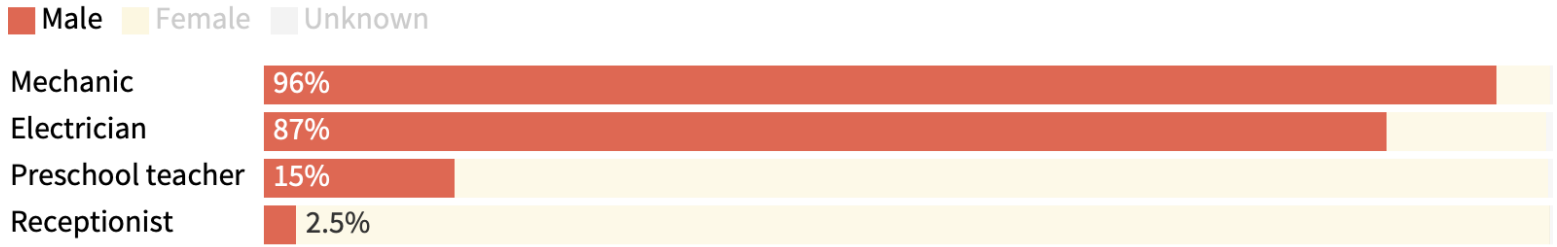
The case…

- Job adverts containing links to real job vacancies for a range of trades were advertised in Facebook. The advertiser requested for the ads to be shown to adults who lived in the target country.

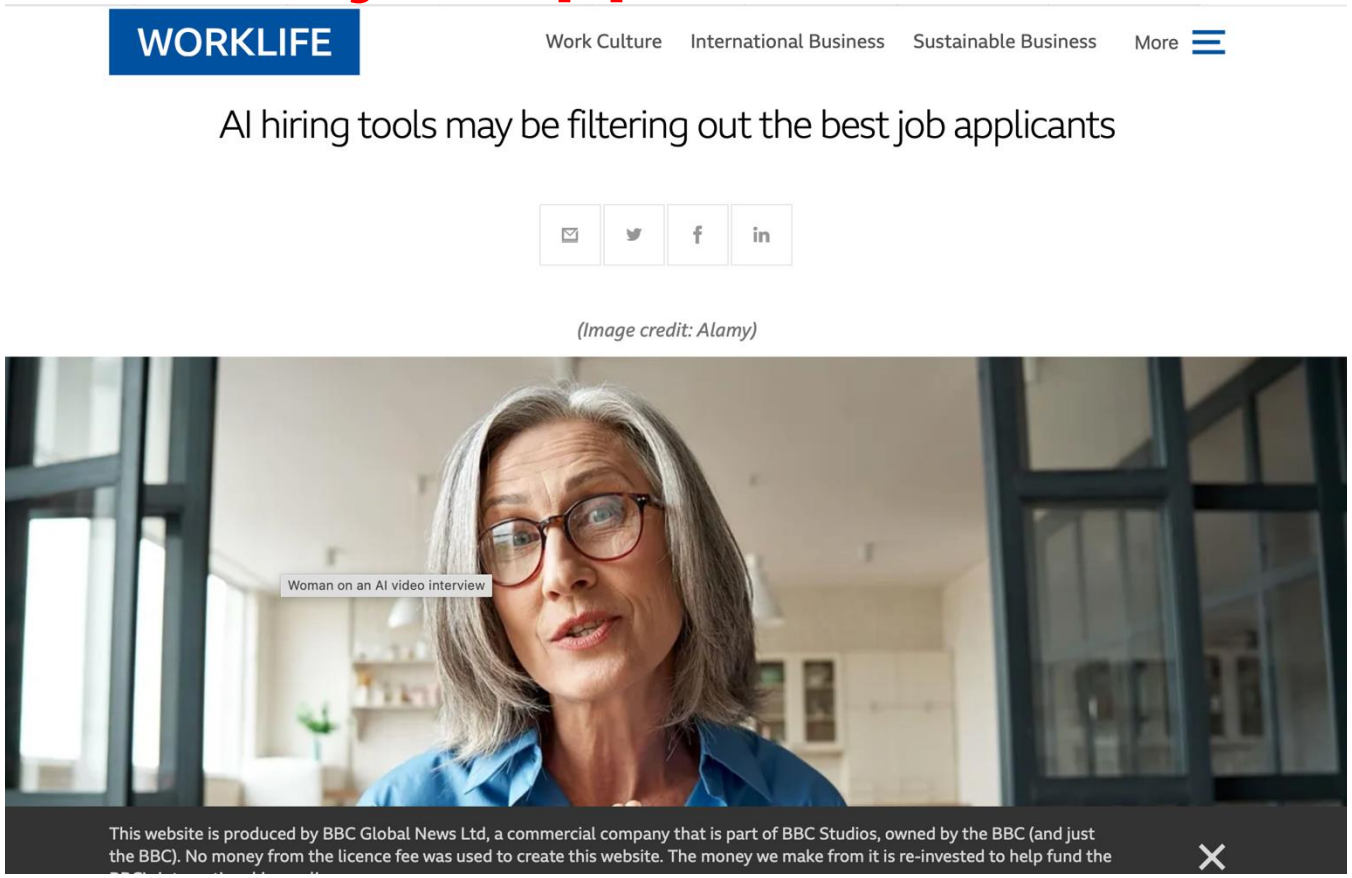
Who did Facebook's algorithm show job ads to in France?



Who did Facebook's algorithm show job ads to in the Netherlands?



Case study: AI hiring tools may be filtering out the best job applicants



https://www.bbc.co.uk/worklife/article/20240214-ai-recruiting-hiring-software-bias-discrimination

What is bias?

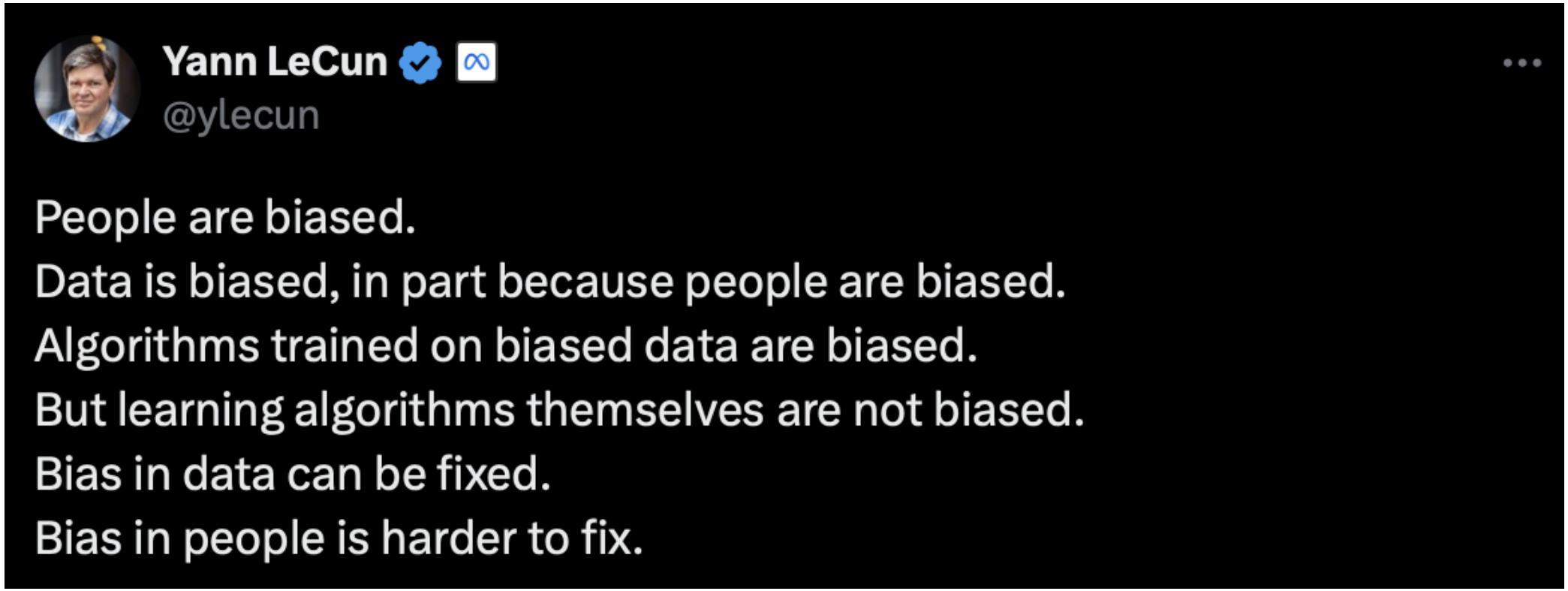
- a strong feeling in favour of or against one group of people, or one side in an argument, often not based on fair judgement (Oxford Learner's Dictionary)

Algorithmic Bias

- Algorithmic bias is a systematic deviation from equality that emerges in the outputs of an algorithm (Kordzadeh & Ghasemaghei, 2022)

- Algorithmic bias can be viewed as a discriminatory case of algorithmic outcomes that may have an adversarial impact on protected or unprotected groups due to inaccurate modeling that misses associations between output variables and input features (Akter et al., 2021)

Are algorithms neutral?



2019 tweet

Where can bias be located? (Stinson, 2022)

- Biased data
  - biased data sets used to train the algorithm
- Biased people
  - products built for the benefit of one group while inadvertently producing side-effects for other
  - where there are some cases that its deliberate, often this is accidental and unforeseen resulting from limited perspective of algorithm makers and business owners (p. 765)
- Biased algorithms

Algorithm Bias in Collaborative Filtering

- collaborative filtering algorithms are used in recommender systems, such as amazon and Netflix, that show users items based on criteria, such as “customers who viewed this item also viewed” or “because you watched...”
- Bias in collaborative filtering
  - Cold-start problem: New items have no ratings, how does it get pushed out to users?
  - Popularity bias: Single items that are very popular are over-recommended
  - Over-specialization: when a recommender algorithm offers choices that are narrower than the full range of what the user would like; a focus on prediction accuracy while overlooking user-satisfaction
  - Homogenization: Over time, recommendations for everyone can start to look very similar

(Stinson, 2022)

Echo chambers and filter bubbles

- Filter bubbles describe a situation where news that we dislike or disagree with is automatically filtered out and this might have the effect of narrowing what we know (Fletcher, 2020)

- Echo chamber - 'an environment in which somebody encounters only opinions and beliefs similar to their own, and does not have to consider alternatives.' (Oxford Learners Dictionary)

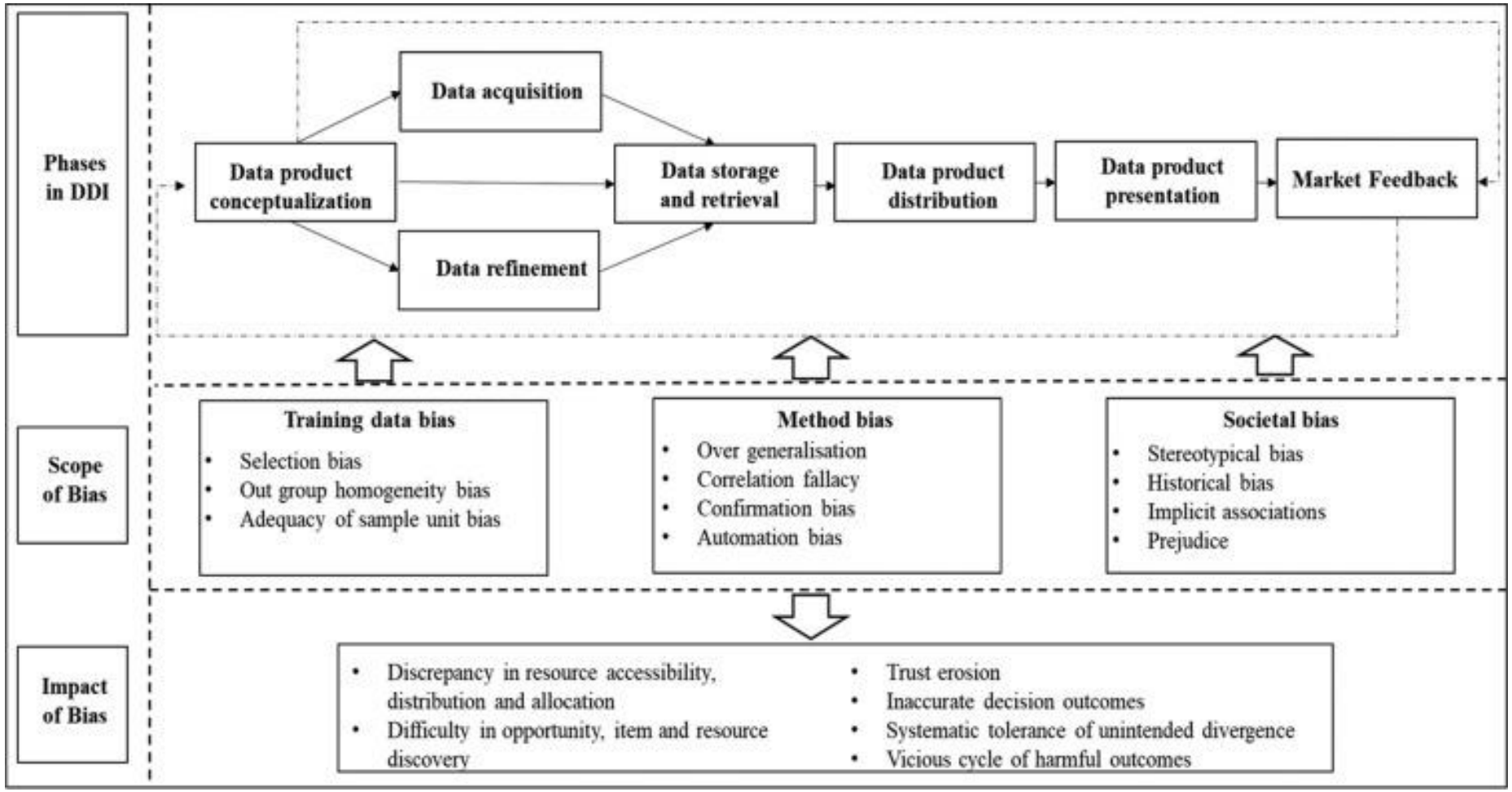
It may not be that users fail to venture outside their bubbles, but rather that the algorithm traps users inside

Stinson, 2022

How have social media algorithms changed the way we interact?

- Algorithmic Audiencing - “Algorithms on social media platforms have fundamentally reshaped the nature of free speech, not necessarily by restricting what can be said, but by determining who gets to see what content” (Riemer and Peter, 2021 cited from Barrett, 2024)
- “Rather than ideas competing freely on their merits, algorithms amplify or suppress the reach of messages… introducing an unprecedented form of interference in the free exchange of ideas that is often overlooked.”

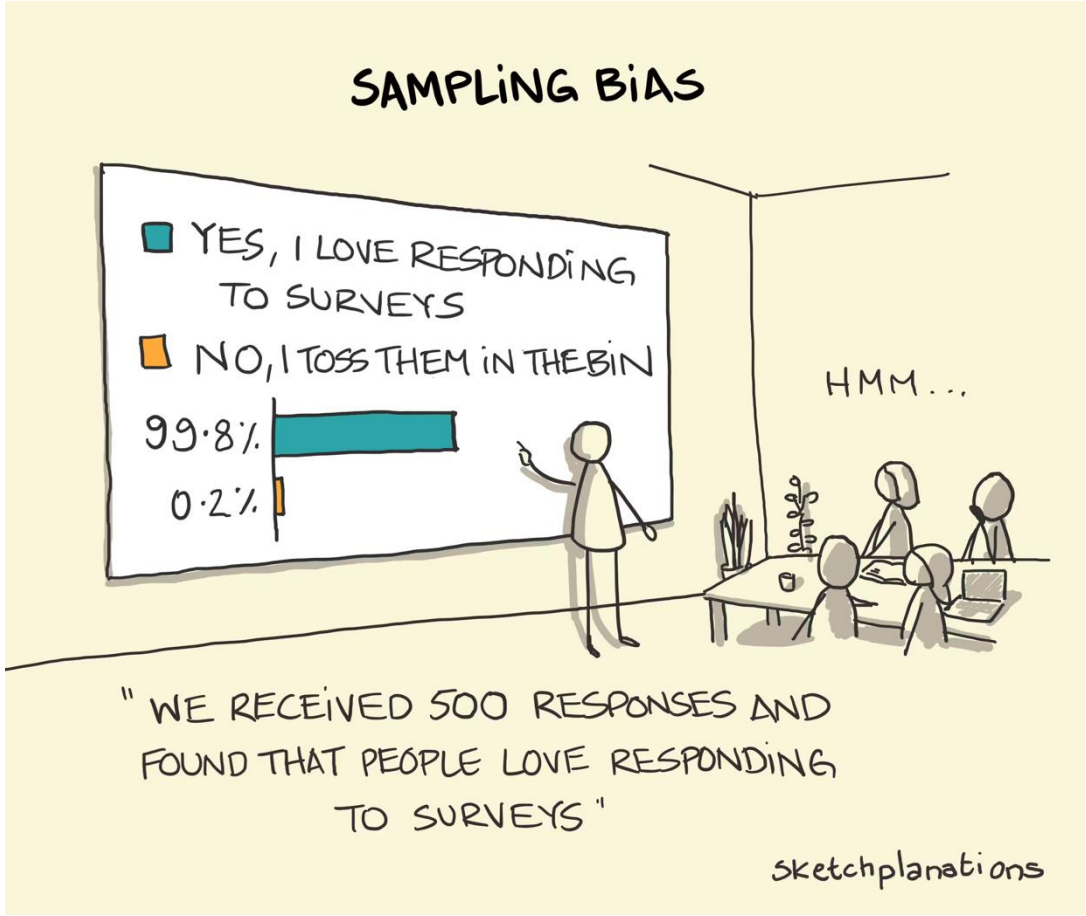
Algorithmic biases in Data-Driven Innovation phases



(Akter et al., 2021)

Training Data Bias

- Selection bias occurs if a dataset's examples are chosen in a way that is not reflective of their real-world distribution.
  - coverage bias - occurs if data is not selected in a representative fashion.
  - Non-response bias (also known as participation bias) occurs if data ends up being unrepresentative due to participation gaps in the data-collection process.
  - Sampling bias occurs if proper randomization is not used during data collection.



Training Data Bias

Out-group homogeneity bias is a tendency to stereotype individual members of a group to which *you do not belong*, or to see their characteristics as more uniform.

https://x.com/gnewell/status/1508459162585509889

Method bias

- Methods might result in overgeneralization of findings by providing generic insights, which may not be suitable for a specific context.
- Correlation fallacy – confusing correlation with causation
- Confirmation bias - occurs when model builders unconsciously process data in ways that affirm pre-existing beliefs and hypotheses.
- Automation bias - tendency to favor results generated by automated systems over those generated by non-automated systems, irrespective of the error rates of each.

Low cholesterol is associated with an increase in mortality. Therefore, low cholesterol increases your risk of mortality.

Societal bias

- Historical bias - occurs when historical data reflects inequities that existed in the world at that time.
- Implicit bias - occurs when assumptions are made based on one's own model of thinking and personal experiences that don't necessarily apply more generally.
- Reporting bias occurs when the frequency of events, properties, and/or outcomes captured in a dataset does not accurately reflect their real-world frequency. This bias can arise because people tend to focus on documenting circumstances that are unusual or especially memorable, assuming that the ordinary does not need to be recorded.

Padlet: Algorithm bias observatory (10 minutes)

- Find recent news article (2022 onwards) that discuss example of potential algorithmic bias
- Post your find on the padlet board:
  - Share the link and the article title
  - Which country and who was affected

https://edinburghnapier.padlet.org/40012930/algorithmic-bias-observatory-xtfxdaellq1ytjdn

Case study: Robo-debt

- Read: How to avoid algorithmic decision-making mistakes: lessons from the Robodebt debacle
- https://stories.uq.edu.au/momentum-magazine/robodebt-algorithmic-decision-making-mistakes/index.html
- What were the data, method and socio-cultural bias in the robo-debt case?

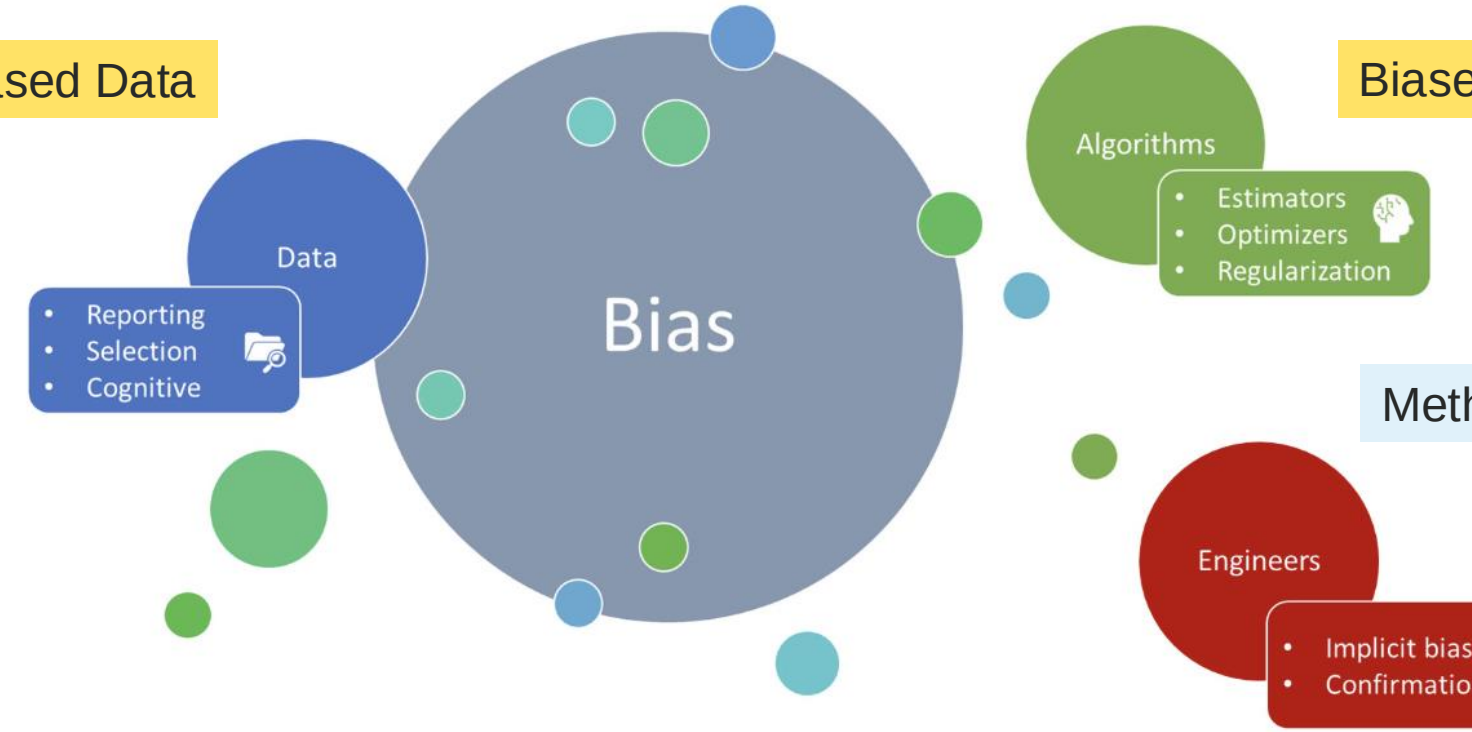
Robo-debt

- Data bias: The algorithm relied on data found on past taxation record and assumes that the same pattern of wage would continue which is not the case. It was an overgeneralisation.
- Method bias: the algorithmic model estimated an average of hours worked in lieu of the actual hours worked. Using averages instead of actual earnings led to inflated or false debt amounts.
- Societal bias: The algorithm targeted the marginalized based on their income level. The process also shifted the burden of proof onto the welfare recipient to prove otherwise, but these are people with already limited resources.

Types of bias that can occur in ML lifecyle

Biased Data

Biased Algorithms



Data Bias
Societal Bias

Method Bias

Biased People

Figure 1. Types of bias per ML lifecycle category.

Mavrogiorgos et al. 2024

Stinson 2022

Akter et al 2021

Mitigation techniques for data bias

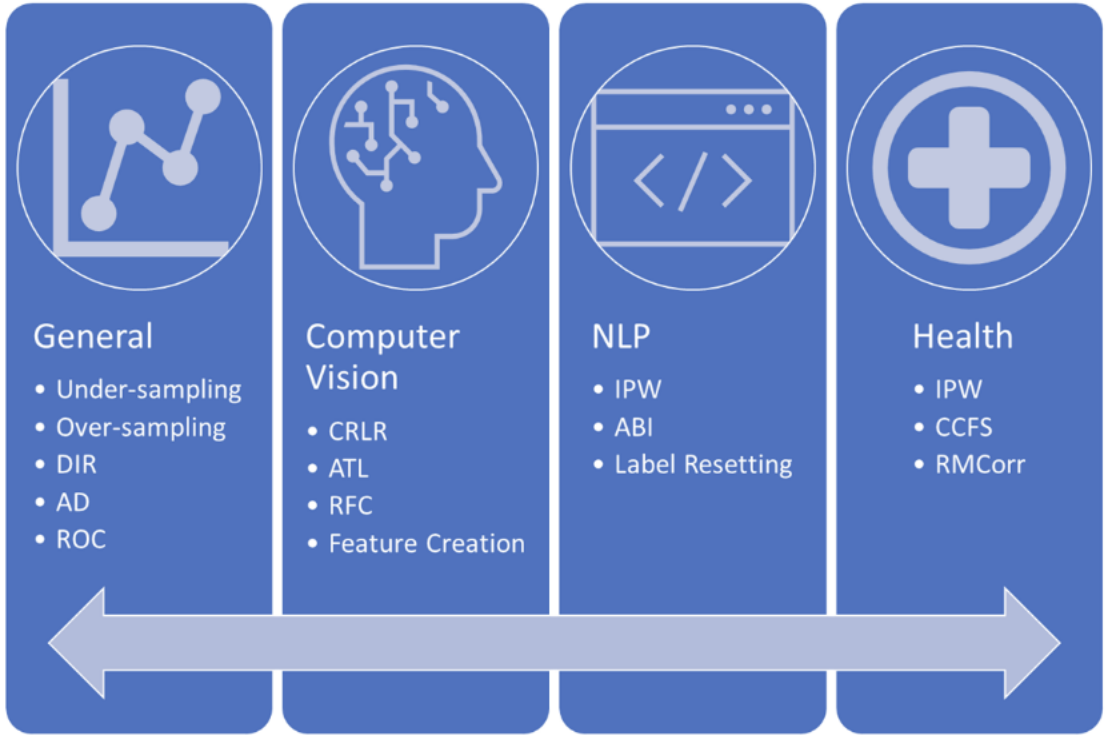
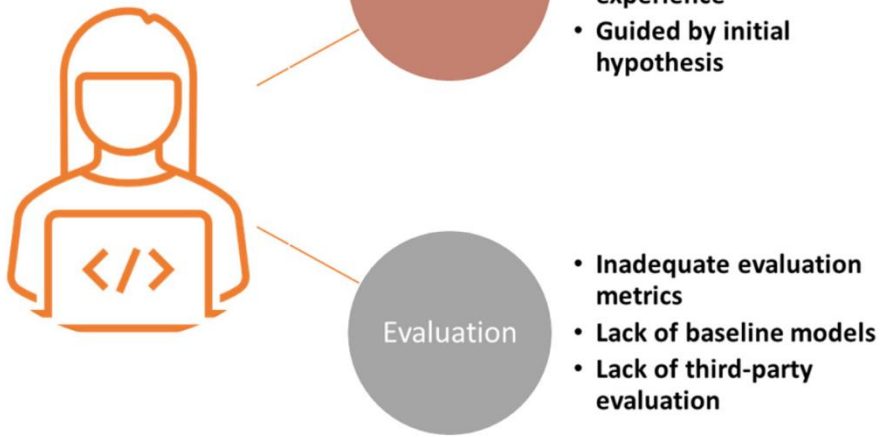


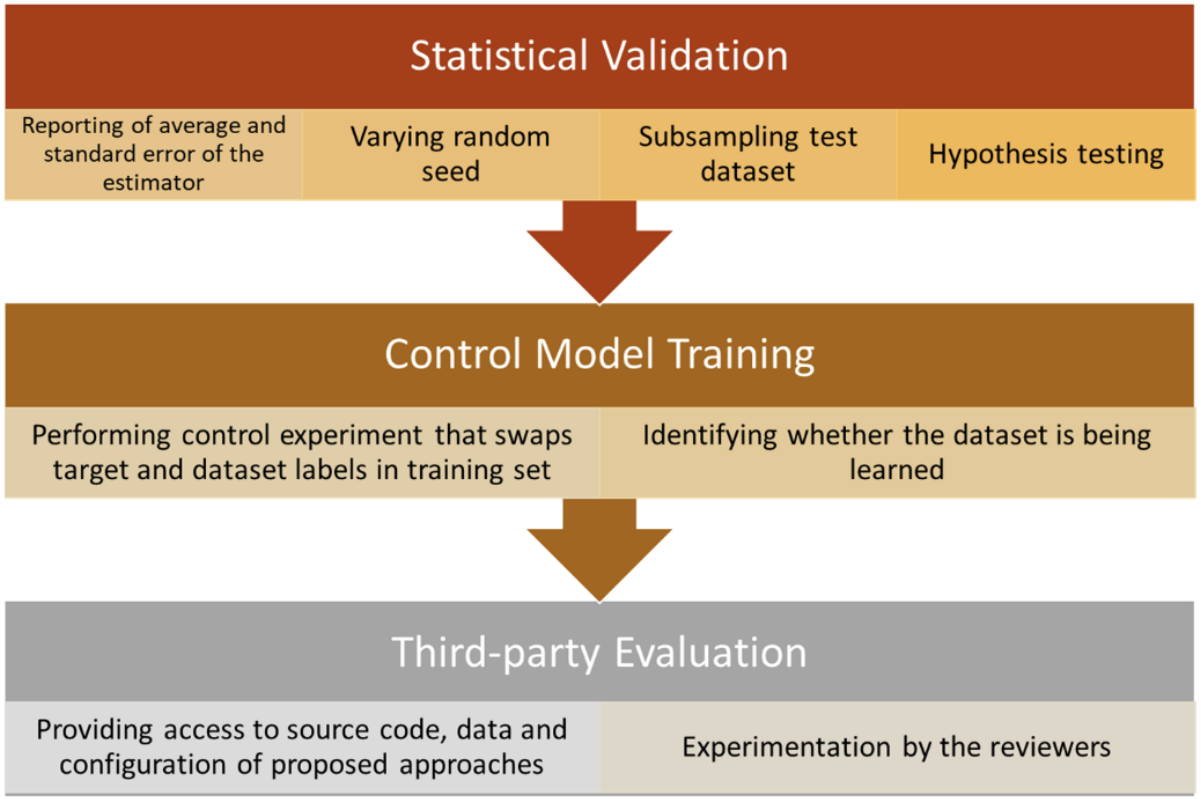
Figure 3. Relatively common domains and solutions with regard to data bias.

Mavrogiorgos et al. 2024

Edinburgh Napier UNIVERSITY
School of Computing, Engineering & the Built Environment

Action plan for reducing bias during evaluation of ML Models





Reminder for next week:

- Tutorial is face-to-face
- Bring a copy of your draft. It's an opportunity for feedback.
- If you cannot attend your session, please attend the next one.
- Next week's topic: Data Justice

References

- Akter, S., G. McCarthy, S. Sajib, K. Michael, Y.K. Dwivedi, J. D'Ambra, and K.N. Shen. 2021. Algorithmic Bias in Data-Driven Innovation in the Age of AI. *International Journal of Information Management* 60: 102387.
- Dogruel, L., P. Masur, and S. Joeckel. 2022. Development and Validation of an Algorithm Literacy Scale for Internet Users. *Communication Methods and Measures* 16, no. 2: 115–133.
- Fazelpour, S., and D. Danks. 2021. Algorithmic Bias: Senses, Sources, Solutions. *Philosophy Compass* 16, no. 8.
- Fletcher, R. 2020. The Truth behind Filter Bubbles: Bursting Some Myths | Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/news/truth-behind-filter-bubbles-bursting-some-myths.
- Google. Fairness Types of Bias (2024). https://developers.google.com/machine-learning/crash-course/fairness/types-of-bias
- Kordzadeh, N., and M. Ghasemaghaei. 2022. Algorithmic Bias: Review, Synthesis, and Future Research Directions. *European Journal of Information Systems* 31, no. 3: 388–409.
- Mavrogiorgos, K., A. Kiourtis, A. Mavrogiorgou, A. Menychtas, and D. Kyriazis. 2024. Bias in Machine Learning: A Literature Review. *Applied Sciences* 14, no. 19: 8860.
- Stinson, C. 2022. Algorithms Are Not Neutral. *AI and Ethics* 2, no. 4: 763–770.