



REVIEW: FAIRNESS



What is Fairness?

- Fairness is a fundamental principle that governs how we treat others in society. It is about making decisions and distributing resources in a way that ensures justice and impartiality.
- Equality means giving everyone the same resources or opportunities, regardless of their circumstances.
- Equity goes a step further, recognizing that people's needs are different, and adjusting resources or opportunities to achieve fair outcomes.

What is Fairness?

CHATGPT Definition : Fairness in AI ethics means ensuring that AI systems make unbiased decisions, treating all individuals and groups equally, and avoiding discrimination based on factors like race, gender, or socioeconomic status. It aims to promote inclusivity and equal opportunity in outcomes.

VS

Philosophical Definition: Treating equals equally and unequals proportionately unequally, within decision domain.

Applying Fairness: Equals

Title: Equal Treatment

- AI systems should not discriminate between individuals with similar qualifications, needs, or circumstances.

Example: Loan approval systems should treat applicants with identical credit scores, income levels, and histories equally.



Applying Fairness: Unequals

- AI should recognize and treat different cases differently if the differences are ethically relevant.
- Example: Healthcare AI should prioritize patients based on severity of illness rather than applying a one-size-fits-all approach.



Historical Perspectives on Fairness

- Fairness has been a central theme in philosophy, ethics, and law for thousands of years, evolving over time based on societal values and needs.
- Aristotle: Introduced the idea of distributive justice, which means allocating resources based on merit or need. He argued that fairness is about proportional equality — treating similar cases similarly and different cases in proportion to their differences.



Modern Views on Fairness

- John Rawls, in his book *A Theory of Justice*, developed the idea of "justice as fairness," advocating for social and economic inequalities to be arranged in ways that benefit the least advantaged in society.
- Today, fairness continues to be debated in the context of social justice, equity, and policies aiming to address systemic inequality.



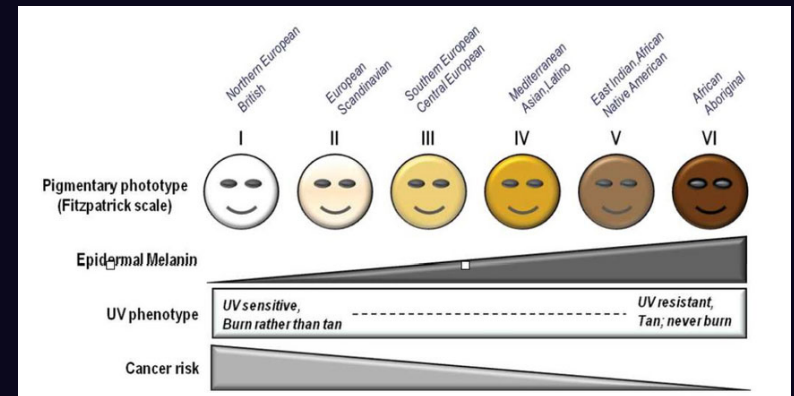
Challenges in Defining Fairness

- Subjectivity: Fairness is often subjective, with different people holding different views on what is fair based on their experiences and values.
- Balancing Equality and Equity: It's challenging to find the right balance between treating everyone equally and adjusting treatment based on individual needs.
- Fairness vs. Efficiency: In business and policy, fairness can conflict with efficiency. For example, equal treatment may seem fair but might overlook individual circumstances, like flexible work arrangements.
- Legal and Moral Dilemmas: Laws may mandate equal treatment, but moral considerations may call for special policies, such as affirmative action, to correct past inequalities.

CASE 1:

FACT: AI Image analyzing for cancer performs better on whiter skin.

Complication: Cancer risks rises with whiter skin.



Discussions:

- Breast cancer is more common in females than in males. Therefore, more resources are allocated to detect breast cancer in females. Is this fair?
- Height is an important factor in basketball. Should a player's height be considered during their selection?

Fairness Within Decision Domains: Examples

- Height in Sports
 - Relevant: Basketball player selection (physical advantage).
 - Irrelevant: Hiring in software engineering (skill-based).
- Financial Status in Loan Approvals
 - Relevant: Financial history and ability to repay.
 - Irrelevant: Race, ethnicity, or other personal attributes.
- Academic Merit in College Admissions
 - Relevant: Test scores, academic performance, extracurriculars.
 - Irrelevant: Political beliefs, personal appearance.

CASE 2: HIRING ALGORITHM DEBATE

- Video Interviews: Algorithms may discriminate based on irrelevant factors, such as:
 - Glasses or facial features
 - Backgrounds (home setting, lighting)

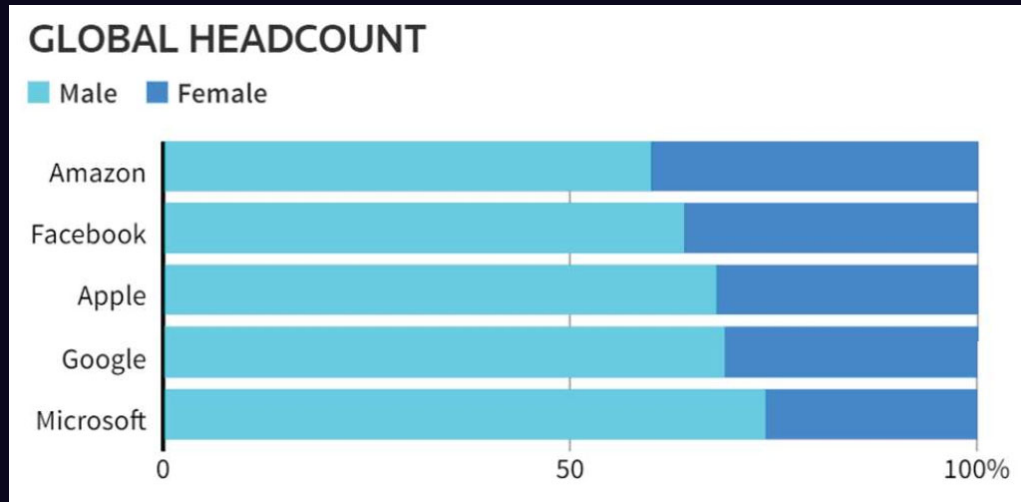
Computer science professor Katharina Zweig:

“The fundamental problem with face recognition by machine learning is that we never know exactly which pattern in an image these machines are responding to.”

CASE 2: HIRING ALGORITHM DEBATE

Amazon's Experimental Hiring Algorithm:

Penalized resumes containing the word “women’s”,
leading to biased evaluations .



Open Discussion: Fairness in Surge Pricing

Non-Relevant Attributes:

- Income Level:
 - Should not disadvantage low-income individuals.
- Geographical Location:
 - Pricing should not vary unfairly by region.

Relevant Attributes:

- Demand Levels:
 - Reflects real-time market needs.
- Service Availability:
 - Encourages providers to meet peak demand.



THANK

YOU