40 Understanding gender bias in AI: A critical reflection exercise

Author:

Javiera Atenas

Contact details:

j.atenas@uos.ac.uk

Role:

Educator

Institution/organisation:

University of Suffolk

Context:

AI is being used to make automated decisions that affect everyday lives, from getting medical treatment to get a job, and women are highly discriminated as the "machine" interprets physical and behavioural data from women in a negative light, missing opportunities to provide with effective care or to get new jobs.

Tool(s) used:

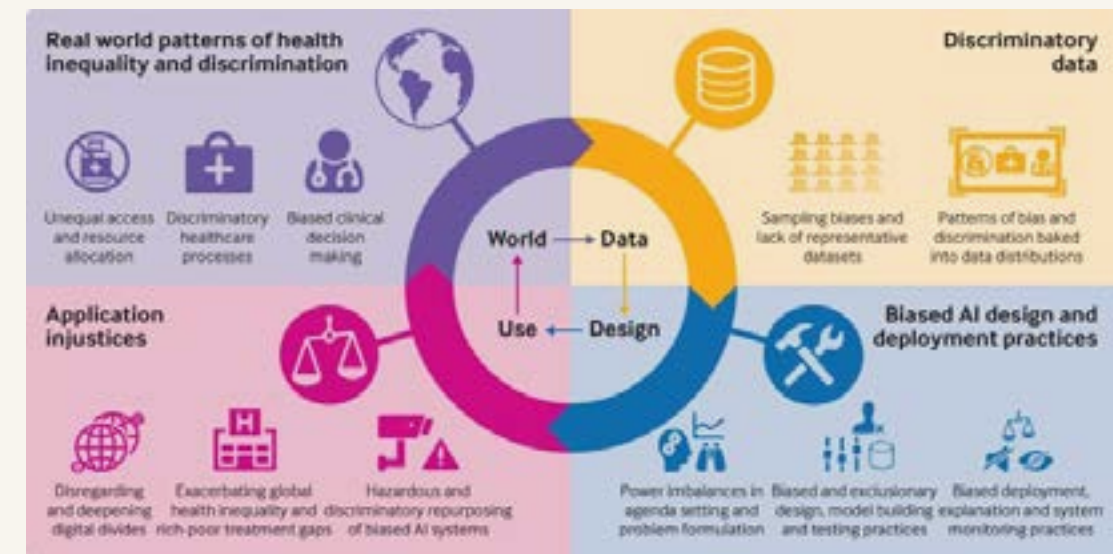ChatGPT, DALL·E and any AI emerging technology

My idea

Explore through AI enabled tools such as ChatGPT and DALL·E how discriminatory data triggers gender bias in AI, prompting questions for text and visual outcomes to then discuss in groups where are the bias and how these can be mitigated

What I aim to achieve

To support educators and students the risks of letting machines to make decisions that can deeply affect the present and future of groups putting them disadvantage widening the gender gap in many different dimensions

Where the inspiration comes from

The data feminism book and current discussions of feminist data leaders and pedagogist about the impact on women and other vulnerable communities.

Infographic of Bias in AI: Discriminatory data; biased AI design and deployment practices, application injustices and real world patterns of health inequality and discrimination.

For more information on Idea 40: ORF's commentary on 'Gender Bias In Futuristic Technologies: A Probe Into AI & Inclusive Solutions and BerkeleyHaas' playbook on 'Mitigating Bias in Artificial Intelligence'.