

We need a global feminist campaign against Artificial Intelligence bias



Ruhi Khan, ESRC Researcher at the LSE, argues that Artificial Intelligence is embedded with patriarchal norms and thus poses many harms to women. She calls for a global feminist campaign to tackle the issue of gender and race bias in AI systems through awareness, inclusion and regulation.

Recently I attended an [event](#) in London on Women in Technology with some brilliant women who occupy prominent position in global tech companies, run start-ups, or study technology. While their success stories are inspiring, every woman in that room knew that the road to gender equality in technology is a long and winding one, fraught with vast difficulties, and the conversation around this has only just begun.

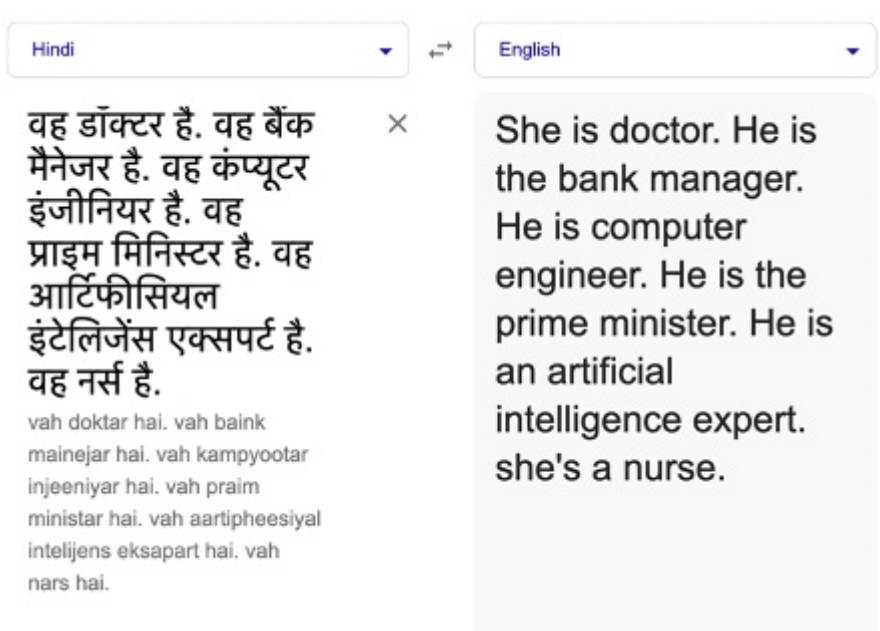
Unsurprisingly, male participation in the discussion was non-existent. Women across race, age, class joined in actively though. The room became a sanctuary where we could tell our stories, pat each other's backs, share concerns, cry over disappointments, point out red flags, hold hands and hope that one day in the distant future the world would be better for women in tech because we were taking that one small step towards promoting gender equality within our own small ecosystems.

Herein lies the problem. And perhaps also the solution.

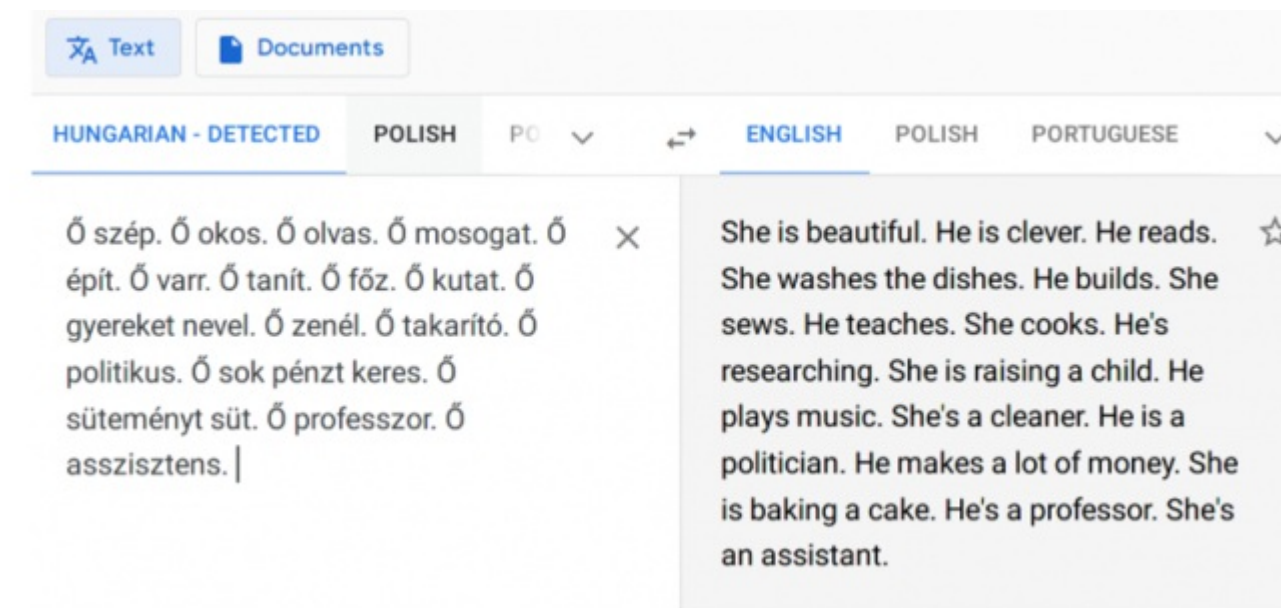
Technology – like education and politics before it – is considered a male bastion and history shows how long and difficult the journey was for women to enter universities and parliament. Yet unlike the right to vote, the [exclusion](#) and the conscious and unconscious bias of technology has not yet led to a concerted campaign for change. There are some incredibly talented [women](#) who are raising their voice against this, but the spark has yet

to catch on and more and more women (and men) need to add their voices, to merge their smaller ecosystems, garner enough attention and demand change on a global stage. The companies that deploy [tech bias](#) are global multi-nationals, so the dissent against them also needs to be a global campaign.

The first step is to acknowledge that the problem exists. Many women stumble across this bias through everyday activities. So here's a simple exercise: Type a few sentences using a gender-neutral pronoun in Google translator in the language you prefer and see what response it throws. I chose Hindi to English, and this is the response I got:



This is not specific to any particular language. For another example, look at [this translation](#) from Hungarian to English. The results here too have patriarchal and misogynist underpinnings. Why do you think this is happening?



Natural language processing (NLP) models might associate certain genders with professions, roles or behaviours based on programmers' bias or skewed patterns in the training data which only reinforces societal stereotypes in technological systems. While a progressive step in the English language vocabulary is to employ a gender-neutral pronoun like 'they' to address an unknown entity rather than force a gender on it, it seems IT systems, while galloping million miles a minute on cutting edge innovations, are still deeply seeped in unfair gender stereotypes.

AI systems are becoming more ubiquitous. A plethora of organisations – both governments and corporations – are racing to harness the power of AI and predictive technologies in hiring, policing, criminal justice, healthcare, product development, marketing and more. These are slowly but surely seeping into the hands of the public through easy-to-use Chat GPT and other generative AI tools that cost nothing or very little. But what is the cost to those who are discriminated against by AI systems and to society as a whole?

AI harms to women

In the automobile industry, the design of seatbelts, headrests and airbags has traditionally been based on data collected from 'male' crash dummies. So, typically female anatomical differences like breast and pregnancy belly were not given any consideration. This made driving far more [unsafe for women](#) with 47% women more

likely to be injured during an accident and 17% more likely to die. Now with self-driving cars, [object recognition](#) and driver monitoring are crucial for vehicles to navigate safely on the road. But if datasets of 'humans' have images of light-skinned people, the machine will not consider dark skinned people as humans, and they will be at increased risk of accidents.

In healthcare, cardio-vascular diseases have long been considered a problem in men and data has been collected from male patients. So even with similar symptoms, men and women would be advised differently by an AI healthcare system. Men would be asked to call emergency services for a suspected heart attack, whereas a woman would be asked to 'de-stress' or seek advice for depression in a few days, even though women also suffer from heart attacks. Bias against underrepresented groups is a huge problem in AI-driven healthcare systems. When it comes to detecting skin cancer, AI tools are not sensitive enough to detect cancer in people with darker skin as they have been developed using images of moles from light skinned individuals. According to one [research](#) study, among 2,436 pictures of lesions where skin colour was stated, only 10 were of brown skin and only one was of dark brown or black skin. Among the 1,585 pictures with information on ethnicity, none were from people with African, Afro-Caribbean, or South Asian backgrounds.

In policing, image recognition systems might misidentify, or label people based on their gender or race or both. This can have serious consequences, especially in applications like security or law enforcement. Facial recognition technology is disproportionately based on data from white (80%) and male (75%) faces, so while the accuracy is 99-100% in detecting white male faces, it falls sharply to 65% for black women. If we deep dive into this [research](#) and compare the statistics of Microsoft and China-based Face++, a rather bleak picture emerges. Accuracy rates for lighter males in Microsoft software is 100% and for lighter females its 98.3% and Face++ also has a remarkable 99.2% accuracy and 94% respectively: But when it comes to darker males and darker females, Microsoft's accuracy rates fall to 94% and 79.2 % and while Face++ is 99.3% accurate for darker males, it falls to a shocking 65.5% for darker females.

In recruitment, AI can determine what postings a candidate can see on job search platforms and decide whether to pass the candidate's details on to a recruiter. It is believed that 70% of companies use automated applicant tracking systems to find and

hire talent. [LinkedIn](#) was taken to court as it paid 686 female employees less than their male counterparts in engineering, marketing and product roles. AI-driven software is likely to mimic historical data that pick men over women. Considering LinkedIn's bias in its own recruitment, it is little wonder that LinkedIn's [algorithm](#) worked against female users by showing them fewer instances of advertisements of high-paying jobs compared to male users even with a similar or better skill set. [Google](#)'s online advertising system displayed high-paying positions to males much more often than women. Offline bias often tends to spill over into online automated systems. [Amazon](#) discovered that its recruitment software was biased against women as it used artificial intelligence to give job candidates scores ranging from one to five stars that were based on observing patterns in resumes submitted to the company over a 10-year period which predominantly came from men. Essentially, the machine learnt from this that men were better than women, and so ranked them much higher.

There is a significant gender gap in data systems, and the design and use of AI systems along with mis-programmed algorithms have dire consequences for women. This gets compounded when we bring in intersectionality. It is important to realise that AI is created by people who can transfer their own biases and those present in society into technology. The selection of training data is also important as a less inclusive data set will impact the machine learning capabilities.

As artificial intelligence becomes a deeper and more intrinsic part of our everyday life, it becomes increasingly important to address the challenges women face due to this bias in a collective and concerted way. Women fought long and hard to get the vote in some countries, the demand for rights for education, employment and abortion among others is an ongoing battle in parts of the world. Now we need a new campaign, one that invites women across the world, of all ages and all colours to question their place in a technologically driven world, to not be silent spectators to the decision of algorithms but active participants in shaping them; to not just accept AI-driven decisions but question their validity and authenticity.

A feminist approach to AI should be three-pronged:

- 1. Awareness**

Bias against women is so deeply entrenched in society that they are often

overlooked and harms against them considered a norm. We must not shy away from addressing problems of sexism and racism. Efforts to mitigate gender bias in AI should be part of a broader initiative to promote fairness, diversity, and inclusion in both technology and society. Campaigns should be based on collective solidarity across countries, race, gender, sex, class, age and other aspects of intersectionality.

2. Inclusion

Diversity is the key to overcoming bias. This includes diversity in terms of people in the tech industry and diversity in the datasets that AI systems are trained on. Women are under-represented in both of these. Few women take up tech-related careers and many stay away from digital services due to circumstances or fear of online harassment. Women's exclusion from the digital sphere has increased the absolute gap between men and women's internet access by [20](#) million since 2019 and has taken £1 trillion off the GDP of low- and middle-income countries. According to [the UN](#), women constitute just 28 per cent of engineering graduates, 22 per cent of artificial intelligence workers, and less than one third of tech sector employees globally. Making STEM education and careers lucrative and accessible for girls would be a step in bridging this gap. Including women in the creation, deployment and audit of AI systems must be mandatory.

3. Regulation

It is essential to have a regulatory framework that can provide assurance that AI is safe and fit to purpose. While some countries have not yet begun a conversation on AI regulations, others are in the process of consolidating it, but the challenge remains that technological advances out-pace legislation. So, a more pro-active and predictive regulatory framework must be sought. A broad participation across various stakeholders is important to distribute the decision-making power to the many and not just leave it in the hands of the very few in the consolidated AI scape. Ethics should be at the heart of any AI policy and its vision to benefits for humanity.

Conclusion

It is time to call upon feminist solidarity to unwrap the black boxes of the complex, convoluted and opaque algorithms, expose the conscious and unconscious bias that exists in the datasets, AI systems and their human developers, to occupy more spaces in the STEM industry, to strategise and innovate tech, to campaign and legislate until AI is ethical and truly inclusive. Women must lead this AI transformation.

This post represents the views of the author and not the position of the Media@LSE blog nor of the London School of Economics and Political Science.

Featured image: Photo by [Markus Spiske](#) on [Unsplash](#)