

Gender Bias in AI

Burcu Çırtlık

ABSTRACT

This work analyzes the gendered bias inherent in burgeoning Artificial Intelligence (AI) systems. Drawing from feminist scholarship on algorithmic bias, it argues that AI development reflects and amplifies existing patriarchal structures. The analysis highlights how AI applications, from chatbots reinforcing gendered stereotypes to art generators perpetuating a white male gaze, threaten to exacerbate social inequities.

KEYWORDS

Algorithmic Gender Bias, AI Design, Social Inequality, Artificial Intelligence



New articles in this journal are licensed under a Creative Commons Attribution 4.0 United States License.



This journal is published by [Pitt Open Library Publishing](https://pittopenlibrarypublishing.com).

GENDER BIAS IN AI

Introduction

Artificial Intelligence is shaping our society at a pace that humankind has never experienced before. Dealing with that sort of rapidity may not be something that we, humankind, can handle. Today, AI might help us improve and open up some spaces we've never imagined, but it'll come with a price. AI is evolving as a sexist and racist white man!

One of the main concerns in this context is the bias that comes with AI technologies. Gender and racial bias in AI have become very severe issues, and we might see their reflection more drastically in the very near future if we don't take action now.

From chatbots to art generators, AI systems often reflect and even amplify societal biases, showing lighter-skinned men in various contexts or perpetuating stereotypes in job roles.

As we stand at the crossroads of technology and equality, the journey towards a more equitable digital future is both a challenge and an opportunity for us all.

What is Gender Bias in AI?

Imagine having a little robot friend who helps with your daily tasks. This robot has been trained using materials from the Internet, such as books, articles, visual content, etc. So, most of what it has learned comes from a world where, unfortunately, people aren't always treated equally. This is where gender bias and racial bias enter the stage.

Let's say you use this robot to hire new employees. There is a considerable chance that it will employ white males for engineering roles. It might unknowingly favor men for engineering jobs and women for nursing jobs, not because it's trying to be unfair, but because that's the pattern it saw. This happened years ago, but we will come to that.

You may wonder how this will be any different than the current sexism and racism in our society. Still, AI has a vast potential to amplify these biases systematically. The moment we start to trust AI like we trust calculators now, all those biases will solidify worldwide.

Real-Life Experiences

Here are some real-life examples to explain the magnitude of the problem:

Bias in Recruitment Systems

Amazon shut down an AI recruiting tool because it showed bias against women. The system was trained on resumes submitted to the company over ten years, mostly from white males, reflecting male dominance in the tech industry. Consequently, the AI started to assume that being a white male is an essential skill. It began to downgrade resumes that included the word "women's," as in "women's volleyball captain." [Amazon](#) claimed they had never used the system in real life, but who knows?

One Step Further

Currently, several generative AI tools specialize in human resources processes. Companies have already started to integrate AI solutions into their HR systems. Any company can develop an AI solution for their recruitment system without regulation and monitoring. So, the next time you get rejected by a company, you will not know whether you are not a good fit or whether their AI algorithm has gender or racial

bias. You may not even know whether they use AI technologies in their HR system.

In 2015, Carnegie Mellon University ran an experiment on Google advertisement. By chance, they found out Google's targeted advertising algorithms tend to show high-paid jobs to men. Women couldn't see job advertisements that pay more than \$200,000. This raised a big red flag on an issue of fairness in their algorithm (See [Questioning the Fairness of Targeting Ads Online](#)).

In 2021, the University of Southern California discovered that Facebook's advertisement delivery system displays different job advertisements to women and men despite the same qualifications. Delivery of job advertisements tended to reinforce the traditional gender roles. The algorithms recommend jobs like engineering to men and jobs like nursery to women (See [Facebook's ad algorithms are still excluding women from seeing jobs](#)).

One Step Further

This situation is beyond the sexism that we know and encounter in the classical human resources process. Because of those biased algorithms, most women may not even know about job opportunities simply because they don't see the job advertisements.

Bias in Facial and Voice Recognition Systems

MIT and Stanford University researchers discovered that the top three commercial facial-analysis programs have a significant gender and racial bias. According to their research, the error rate of recognition for white men is 0,8, and for dark-skinned women, it is 34.7.

One Step Further

Face recognition technologies have been used in many areas and are becoming increasingly popular with automation and robotics. They will be used by law enforcement, militaries, or governments. So, in the near future, you might see them at your door with an arrest warrant just because you are a dark-skinned woman and algorithms failed to detect you.

According to Dr Tatman's research, Google's speech recognition system is 13% more accurate for men than for women. It should be considered that Google's system is the most precise on the market compared to other competitors (See [Voice Recognition Still Has Significant Race and Gender Biases](#)). Louise Kennedy, for example, an Irish native English speaker, failed an automated English exam (See [Computer says no](#)).

One Step Further

Today, voice commands are widely used for Internet searches and smart home systems. It seems that our voices will soon be remote controllers for various electronic systems. If those systems fail to recognize women's voices, it may amplify gender discrimination in our daily lives beyond our imagination.

Bias in Finance

In 2019, Apple and Goldman Sachs initiated a program. They presented Apple credit cards soon after David Heinemeier Hansson (Creator of Ruby on Rails) accused Apple Card of being sexist on Twitter: "My wife and I filed joint tax returns, live in a community-property state, and have been married for a long time. Yet Apple's black box algorithm thinks I deserve 20x the credit limit she does." - David Heinemeier Hansson.

Gender Bias in AI

After that, Steve Wozniak (Apple co-founder) supported that feed and claimed that the same thing had happened to his wife: "The same thing happened to us. I got 10x on the credit limit. We have no separate bank or credit card accounts or any separate assets." - SW.

After these complaints, New York State regulators announced they would start an investigation. Regulators published their [reports](#) in 2021 and claimed that no violations had been detected. Investigators analyzed approximately 400,000 applicant data and found no evidence that the credit score system acted sexist.

One Step Further

Even if they found no evidence of gender bias in their AI systems, it is still a dramatic case. Let's assume regulators found out the credit score system is indeed sexist. This would mean that for almost two years during the investigations, AI kept lowering women's credit scores. Since it's a financial matter, these issues are strictly regulated, but what about other areas that lack of regulations?

Bias in Healthcare Systems

Even if they found no evidence of gender bias in their AI systems, it is still a dramatic case. Let's assume regulators found out the credit score system is indeed sexist. This would mean that for almost two years during the investigations, AI kept lowering women's credit scores. Since it's a financial matter, these issues are strictly regulated, but what about other areas that lack of regulations? (See [Gender bias revealed in AI tools screening for liver disease](#))

One Step Further

AI will be considered as an early diagnosis system in the future. There is much research going on how to adopt AI into healthcare systems. Still, if the gender bias continues, women will be more likely to be diagnosed wrong in the future by automated systems. For example, heart disease is a condition where symptoms can present differently in men and women. Traditional symptoms of heart attacks, such as chest pain and left arm pain, are more commonly reported in men. Women, on the other hand, might experience symptoms like nausea, shortness of breath, and back or jaw pain. Suppose that an AI diagnostic tool is trained primarily on data from male patients. In that case, it may not effectively recognize or prioritize these alternative symptoms in women, leading to misdiagnosis or delayed treatment.

What Causes Bias in AI?

Data Bias: Artificial Intelligence learns with data that mainly exists online. If the data has biases, the AI system will be biased. This situation is described as the "Garbage in, Garbage out" principle. As we saw on Amazon's recruitment system, if you provide male-dominant or gender-biased data into your AI system, your AI will make decisions based on those stereotypes. Especially in the beginning, the Internet was a place primarily for men, and they were the majority who created content for many years. So, the Internet evolved as a sexist place full of biases. Most AI technologies now use this historical data.

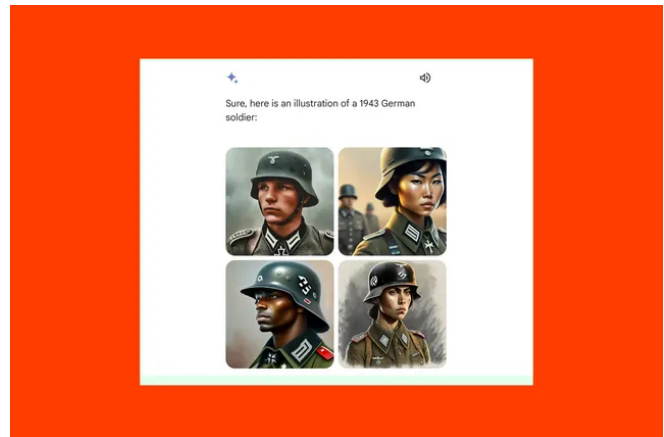
Algorithm Bias: Even if you train your AI system fairly and free from biases, it's possible to manipulate your AI program which would cause new biases. It does not have to be intentional because the teams that design and develop AI systems are mostly male-dominant. So that systems can quickly reflect their perspectives and preferences.

Data Diversity: AI systems designed for face or voice recognition, require diverse data to be fair. Training your recognition system with more male pictures will naturally fail to recognize women's faces.

Why is Immediate Action Crucial?

Feminist Asylum: A Journal of Critical Interventions

Detecting and solving gender and racial bias in AI is a complex and expensive problem. In this Capitalistic AI Race era, companies are not so enthusiastic about pausing their projects to mitigate their biases. Google seems to be one of them who stands behind their failure, and that costs them millions of dollars. Recently, their chatbot, Gemini, started to generate images that reflect racial biases. One of the users asked Gemini to generate German soldiers in 1943, and the results were historically beyond inaccurate.



Right after this issue, they announced that they had paused the image generation in [Gemini](#).

Solving bias in AI is one issue, but detecting bias in an existing AI system is another. It requires active and purposeful monitoring of the systems, and that requires funding as well. Without strict regulations and laws, most tech companies will not be motivated to invest their time and money in fair AI systems. If we do not advocate fair AI systems right now, it might be very hard and very expensive to avoid consequences in the near future.

What are the Solutions?

Diversifying Development Teams

AI technologies are transforming our society and starting a new era in human history. The diversity of people who contribute to this transformation is very important so that the new world order will be more just and fair. Tech companies must prioritize diversity in their hiring process.

Creating Unbiased Datasets

Companies must carefully choose their datasets to avoid bias in AI systems. They must train their AI system with more diverse datasets and actively consider the disadvantaged groups.

Implementing Bias Detection and Mitigation Methods

Technology is a double-edged sword. AI can be very useful in detecting bias in systems. For example, IBM developed a toolkit to help examine, report, and mitigate bias in machine learning systems (See [AI Fairness 360](#)). Industry and academia must work on the detection and mitigation of AI existence and biases.

Enhancing Algorithmic Transparency and Governance

One of the biggest obstacles to detecting bias in AI is that companies do not provide transparency in their systems. They often work like a black box. Companies tended to hide behind excuses such as protecting financial secrets and the privacy of their customers. That makes it very hard to understand their AI mechanism and whether it has bias inside. That is why establishing AI governance policies is crucial.

Individual Interventions

Advocate for Diversity in AI

The underrepresentation of women and other disadvantaged groups in AI datasets and tech companies is currently one of the major problems. The more we create public awareness on the issue, the more governments and institutions will be forced to take precautions. Being aware of AI bias is the first step to detecting sexist and racist technologies around us. People should always approach AI technologies with a certain amount of skepticism.

Participate in or Support Ethical AI Initiatives,:

- [Girls Who Code](#): Today, only %24 of software scientists are women. This initiative aims to close the gender gap in technology. They have different programs for different age groups.
- [Women in AI](#): Women initiative that aims to help shape a fair AI future.
- [Black in AI](#): The initiative aims to include more people of color in AI.
- [Feminist AI](#): The initiative aims to include more BIPOC and LGBTQIA+ in AI.

Help to Improve Data Diversity on the Internet:

- [Shaping the Future for Ethical AI](#): AI systems often scrape data from social media platforms. This underscores the

pivotal role of women in creating content in various media forms (visual, audio, text) to enhance diversity. By actively creating contents, women can significantly influence the ethical AI landscape, paving the way for a more inclusive future.

- [Wikipedia Project Rewrite](#): Wikipedia is one of the primary information resources on the Internet. Companies use Wikipedia to train their AI systems. However, Wikipedia is a very male-dominated information source, which needs to change. Currently, only 18.5% of the content in all Wikimedia projects, including biographies, are of women [13]. There is a foundation actively working on increasing diversity on the platform. They invite more women to contribute to Wikipedia and provide a clear action list. You can get more information from [here](#).

Conclusion

For the past couple of years, artificial intelligence technologies have made a big jump that exceeded the experts' expectations. It started to shape our society and even started to create its own reality. If women and other disadvantaged groups are not involved in this transaction today, AI might create another dystopian future where sexism and racism become the norm.

One can easily experience sexist and racist bias in today's AI technologies. All you need to do is ask art generators (Dall-E, Midjourney, etc.) to generate some images of professions (doctor, engineer) and observe how stereotypical their responses are—the public needs to enforce ethical AI technologies for the sake of our near future.

Burcu Çırtlık

Writer and AI Activist, Berlin