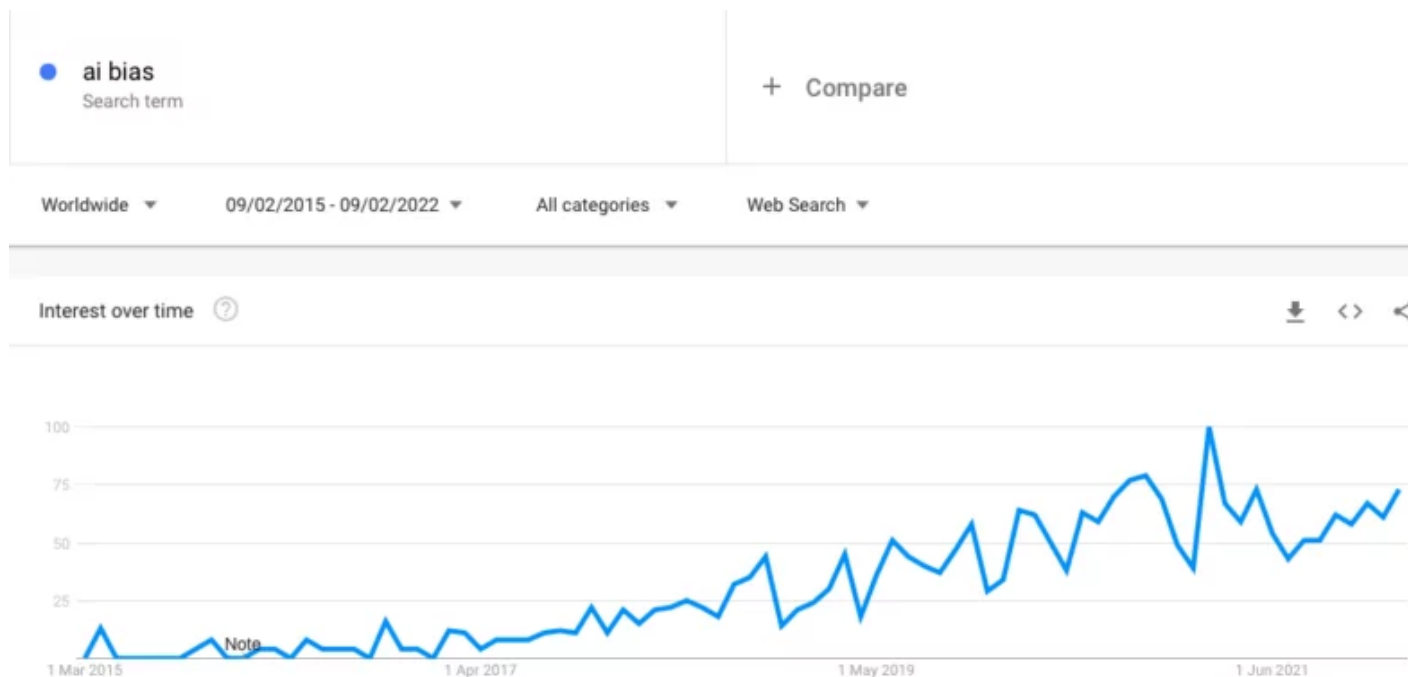


Bias in AI: What it is, Types, Examples & 6 Ways to Fix it in 2022

UPDATED ON MAY 12, 2022 | PUBLISHED ON SEPTEMBER 12, 2020 | 7 MINUTE READ



Written by Cem Dilmegani



Interest in Artificial Intelligence (AI) is increasing as more individuals and businesses witness its benefits in various use cases. However, there are also some valid concerns surrounding AI technology.

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

- Will AI be a threat to our jobs? Yes, 44% of low education workers will be at risk of technological unemployment by 2030.
- Can we trust the judgment of AI systems? Not yet, AI technology may inherit human biases due to biases in training data

In this article, we focus on AI bias and will answer all important questions regarding biases in artificial intelligence algorithms from types and examples of AI biases to removing those bias from AI algorithms.

What is AI bias?



AI bias is an anomaly in the output of machine learning algorithms, due to the prejudiced assumptions made during the algorithm development process or prejudices in the training data.

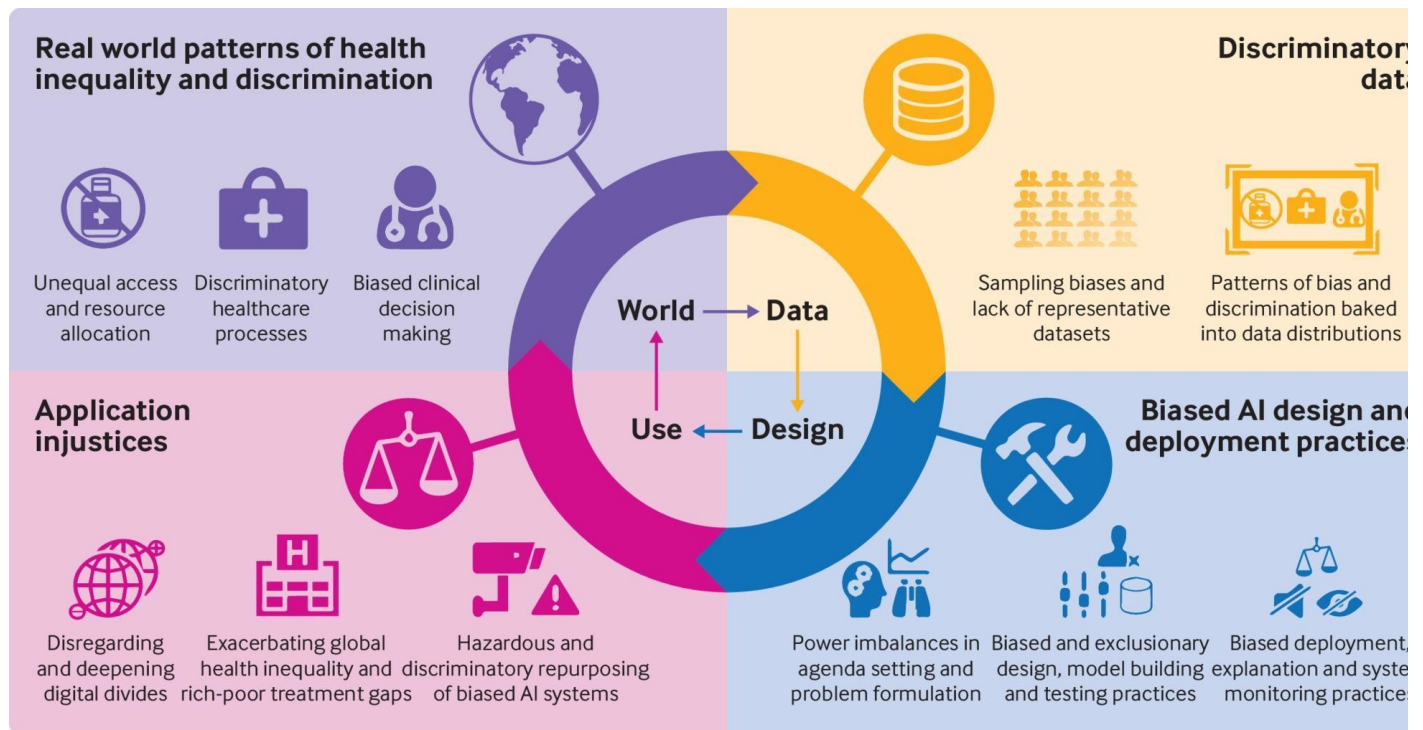
What are the types of AI bias?

AI systems contain biases due to two reasons:

- **Cognitive biases:** These are unconscious errors in thinking that affects individuals' judgements and decisions. These biases arise from the brain's attempt to simplify processing information about the world. More than 180 human biases have been define and classified by psychologists. Cognitive biases could seep into machine learning algorithms via either
 - o designers unknowingly introducing them to the model

undergraduate students which are a specific group and do not represent the whole population.

Figure 1. Inequality and discrimination in the design and use of AI in healthcare applications



Source: British Medical Journal

Will AI ever be completely unbiased?

Technically, yes. An AI system can be as good as the quality of its input data. If you can clean your training dataset from conscious and unconscious assumptions on race, gender, or other ideological concepts, you are able to build an AI system that makes unbiased data-driven decisions.

However, in the real world, we don't expect AI to ever be completely unbiased any time soon due to the same argument we provided above. AI can be as good as data and people are the ones who create data. There are numerous human biases and ongoing identification of new biases is increasing the total number constantly. Therefore, it may not be possible to have a completely unbiased human mind so does AI system. After all, humans are creating the biases.

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

What we can do about AI bias is to minimize it by testing data and algorithms and developing AI systems with responsible AI principles in mind.

How to fix biases in AI and machine learning algorithms?

Firstly, if your data set is complete, you should acknowledge that AI biases can only happen due to the prejudices of humankind and you should focus on removing those prejudices from the data set. However, it is not as easy as it sounds.

A naive approach is removing protected classes (such as sex or race) from data and deleting the labels that make the algorithm biased. Yet, this approach may not work because removed labels may affect the understanding of the model and your results' accuracy may get worse.

So there are no quick fixes to removing all biases but there are high level recommendations from consultants like McKinsey highlighting the best practices of AI bias minimization:

Minimizing bias will be critical if artificial intelligence is to reach its potential and increase people's trust in the systems.

Six potential ways forward for artificial-intelligence (AI) practitioners and business and policy leaders to consider

1	2	3	4	5	6
					
Be aware of contexts in which AI can help correct for bias and those in which there is high risk for AI to exacerbate bias	Establish processes and practices to test for and mitigate bias in AI systems	Engage in fact-based conversations about potential biases in human decisions	Fully explore how humans and machines can best work together	Invest more in bias research, make more data available for research (while respecting privacy), and adopt a multidisciplinary approach	Invest more in diversifying the AI field itself

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

Steps to fixing bias in AI systems:

1. **Fathom the algorithm and data** to assess where the risk of unfairness is high. For instance:
 - **Examine the training dataset** for whether it is representative and large enough to prevent common biases such as sampling bias.
 - **Conduct subpopulation analysis** that involves calculating model metrics for specific groups in the dataset. This can help determine if the model performance is identical across subpopulations.
 - **Monitor the model over time** against biases. The outcome of ML algorithms can change as they learn or as training data changes.
2. **Establish a debiasing strategy** within your overall AI strategy that contains a portfolio of technical, operational and organizational actions:
 - **Technical strategy** involves tools that can help you identify potential sources of bias and reveal the traits in the data that affects the accuracy of the model
 - **Operational strategies** include improving data collection processes using internal “red teams” and third party auditors. You can find more practices from [Google AI's research on fairness](#)
 - **Organizational strategy** includes establishing a workplace where metrics and processes are transparently presented
3. **Improve human-driven processes as you identify biases in training data.** Model building and evaluation can highlight biases that have gone unnoticed for a long time. In the process of building AI models, companies can identify these biases and use this knowledge to understand the reasons for bias. Through training, process design and cultural changes, companies can improve the actual process to reduce bias.
4. **Decide on use cases where automated decision making should be preferred** and when humans should be involved.
5. **Follow a multidisciplinary approach.** Research and development are key to minimizing the bias in data sets and algorithms. Eliminating bias is a multidisciplinary strategy that consists of ethicists, social scientists, and experts who best understand the nuances of each application area in the process. Therefore, companies should seek to include such experts in their AI projects.
6. **Diversify your organisation.** Diversity in the AI community eases the identification of biases. People that first notice bias issues are mostly users who are from that specific minority community. Therefore, maintaining a diverse AI team can help you mitigate

IBM released an open-source library to detect and mitigate biases in unsupervised learning algorithms that currently has 34 contributors (as of September 2020) on Github. The library called [AI Fairness 360](#) and it enables AI programmers to

- test biases in models and datasets with a comprehensive set of metrics.
- mitigate biases with the help of 12 packaged algorithms such as Learning Fair Representations, Reject Option Classification, Disparate Impact Remover.

However, AI Fairness 360's bias detection and mitigation algorithms are designed for binary classification problems that's why it needs to be extended to multiclass and regression problems if your problem is more complex.

IBM Watson OpenScale

IBM's [Watson OpenScale](#) performs bias checking and mitigation in real time when AI is making its decisions.

Google's What-If Tool

Using [What-If Tool](#), you can test performance in hypothetical situations, analyze the importance of different data features, and visualize model behavior across multiple models and subsets of input data, and for different ML fairness metrics.

What are some examples of AI bias?

Amazon's biased recruiting tool

With the dream of [automating the recruiting process](#), Amazon started an AI project in 2014. Their project was solely based on reviewing job applicants' resumes and rating applicants by using AI-powered algorithms so that recruiters don't spend time on manual resume screen

Amazon had used historical data from the last 10-years to train their AI model. Historical data contained biases against women since there was a male dominance across the tech industry and men were forming 60% of Amazon's employees. Therefore Amazon's recruiting system incorrectly learnt that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." Therefore, Amazon stopped using the algorithm for recruiting purposes.

Amazon's Sexist Recruitment AI



Racial bias in healthcare risk algorithm

A health care risk-prediction algorithm that is used on more than 200 million U.S. citizens, demonstrated racial bias because it relied on a faulty metric for determining the need.

The algorithm was designed to predict which patients would likely need extra medical care, however, then it is revealed that the algorithm was producing faulty results that favor white patients over black patients.

Bias in Facebook ads

There are numerous examples of human bias and we see that happening in tech platforms. Since data on tech platforms is later used to train machine learning models, these biases lead to biased machine learning models.

In 2019, Facebook was allowing its advertisers to intentionally target adverts according to gender, race, and religion. For instance, women were prioritized in job adverts for roles in nursing or secretarial work, whereas job ads for janitors and taxi drivers had been mostly shown to men, in particular men from minority backgrounds.

As a result, Facebook will no longer allow employers to specify age, gender or race targeting in its ads.

Extra resources

Krita Sharma's Ted Talk

Krita Sharma, who is an artificial intelligence technologist and business executive, is explaining how the lack of diversity in tech is creeping into AI and is providing three ways to make more ethical algorithms:

How to keep human bias out of AI | Kriti Sharma



We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

Barak Turovsky at 2020 Shelly Palmer Innovation Series Summit

Barak Turovsky, who is the product director at Google AI, is explaining how Google Translate dealing with AI bias:

AI Bias Explained by Google at the Shelly Palmer Innovation Series Summit at ...



Hope this clarifies some of the major points regarding biases in AI. For more on how AI is changing the world, you can check out articles on [AI](#), [AI technologies](#) and AI applications in [marketing](#), [sales](#), [customer service](#), [IT](#), [data](#) or [analytics](#).

Also, feel free to [follow our Linkedin page](#) where we share how AI is [impacting businesses and individuals](#) or our [Twitter account](#).

If you are looking for AI vendors, you can benefit from our data-driven lists of:

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

If you have a business problem that is not addressed here:

IDENTIFY PARTNERS TO BUILD CUSTOM AI SOLUTIONS

HOW USEFUL WAS THIS POST?

CLICK ON A STAR TO RATE IT!



[in](#) Share

[Tweet](#)



Cem Dilmegani

Cem founded the high tech industry analyst AIMultiple in 2017. AIMultiple informs ~1M businesses (as per similarWeb) including 55% of Fortune 500 every month.

Throughout his career, Cem served as a tech consultant, tech buyer and tech entrepreneur. He advised enterprises on their technology decisions at McKinsey & Company and Altman Solon for more than a decade. He led technology strategy and procurement of a telco while reporting to the CEO. He has also led commercial growth of deep tech companies that reached from 0 to 3M annual recurring revenue within 2 years.

Cem regularly speaks at international technology conferences. He graduated from Bogazici University as a computer engineer and holds an MBA from Columbia Business School.

[in](#)

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

Related Posts



HEALTHCARE,
ARTIFICIAL INTELLIGENCE

AI in X-ray Analysis: Benefits
and Challenges in 2022



ARTIFICIAL INTELLIGENCE

6 Components of a Successful
Enterprise AI Strategy in 2022



SUPPLY CHAIN & LOGISTICS,
ARTIFICIAL INTELLIGENCE

Top 5 AI Use Cases for Supply
Chain Optimization in 2022

Leave a Reply

YOUR EMAIL ADDRESS WILL NOT BE PUBLISHED. REQUIRED FIELDS ARE MARKED *

Comment *

Name *

Email *

Post Comment

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

IN THIS POST

What is AI bias?

What are the types of AI bias?

Will AI ever be completely unbiased?

How to fix biases in AI and machine learning algorithms?

What are some examples of AI bias?

Extra resources

 SHARE

 TWEET

Subscribe to the latest news & updates from our experts

Enter your email

Subscribe

☐ By checking this box, you confirm that you have read and agreed to our terms and conditions.



Solutions

RPA

Data Annotation

Process Mining

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking "Accept", you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)

For Tech Users

[Learn Use Cases](#)
[Shortlist Solutions](#)
[Get Advice](#)

Vendors

[Claim & Grow Your Product](#)
[Learn Best Practices](#)

Investors

[Identify Hidden Gems](#)
[Tech Firms By Country](#)
[Tech Firms By City](#)

AIMultiple

[Mission](#)
[About](#)
[Career](#)
[Contact](#)

Data-driven, Transparent, Practical New Tech Industry Analysis

© Copyright 2022 AIMultiple | [Terms of Use](#) | [Privacy Policy](#)

This site is protected by reCAPTCHA and the Google Privacy Policy and Terms of Service apply.

We use cookies on our website to give you the most relevant experience by remembering your preferences and repeat visits. By clicking “Accept”, you consent to the use of ALL the cookies.

[Cookie settings](#)

[ACCEPT](#)