



ARE AI SYSTEMS NEUTRAL?

„Although you answered everything correctly in the query, you won't get an A because I don't like the color of your pants!“

If your teacher were to grade you like that, you would surely find it quite unfair, wouldn't you? Unfortunately, in our society, many situations occur where important decisions are made this way: someone is denied an apartment because of her or his skin color, a foreign-sounding surname or gender leads to rejection when applying for a super cool internship, or someone is avoided because of their religion. **Why do people behave like this? We have prejudices that we have built up through our environment.** We rely solely on our feelings without even knowing if our assumptions are actually true. Many also fear anything that is unfamiliar or different from what they are used to. Therefore, they behave defensively and judge other people based on traits that say nothing about a person's character, behavior, or attitude – especially nothing bad! The color of your pants also says nothing about you as a person!

Discrimination based on origin, skin color, gender, religion, sexual orientation, physical impairment or many other factors occurs constantly in society. We can only fight against this if we all collectively resist it, educate each other about it and support each other. Fortunately, there are also many organizations that uncover and combat such injustices. Nevertheless, we haven't achieved our goal of fair and just treatment for everyone in society yet.



Mrs. Sara Myers

Very well suited. Impeccable and promising curriculum vitae.

Bachelor of Arts (Industrial Information Systems): 1.3
Master of Arts (Industrial Information Systems): 1.3

Several years of professional experience in software engineering.



Mr. Tom Smith

Does not meet all requirements. Qualification questionable.

Bachelor of Science (Mathematics) in Harvard: 3.4
Master of Science (Mathematics) in Harvard: 3.2

No professional experience.
Frequent change of employer.

Rejected due to:

Discrimination!

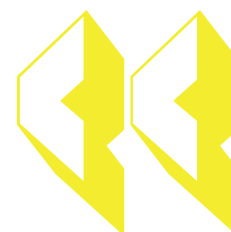


**FAIR MACHINES?
NO WAY!**

If people are so biased, wouldn't it be more sensible to leave important decisions to AI systems? They have no feelings and can therefore judge completely neutrally and without prejudice – that would be pretty great, wouldn't it? Unfortunately, it does not work like that in reality: As you may have seen at other stations, AI systems need to learn their behavior first. And for that, they need a lot of data. For example, if an AI system is supposed to decide which applicant gets an apartment or an apprenticeship, then the AI system must first learn, based on data, who has previously got an apartment or an apprenticeship position. And where do these data come from? Right, from human decisions! And as you saw in the text above, these can be quite unfair.

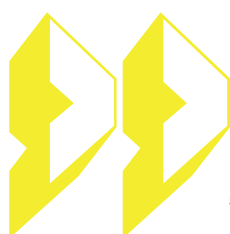
AI systems that learn their behavior from data generated and influenced by humans learn the same behaviors that we humans show. **So, they also have biases in their decisions and discriminate against certain population groups.** You may have noticed this when trying out the station. There are also real-world examples of discriminatory AI systems. For example, in 2015, there was a racism scandal involving Google Photos' image recognition, as the system marked people with darker skin tones in photos as gorillas. This is not only wrong but also very offensive, and it was because Google's algorithm apparently learned the appearance of people only from images of lighter-skinned individuals. Thus, the training data did not accurately represent society and was biased.

Bias usually arises due to the data used, but it can also be caused by incorrect statistical measurements or errors in the algorithm. You probably would not want to use an AI system with bias, would you? With people making decisions, at least you can discuss their decisions and discuss or refute their prejudices. Machines cannot do that.





If all this is so bad, why are there AI systems making decisions at all? By knowing that many decisions have bias, we can also try to combat this in AI systems! To do this, attempts are made to delete from the data all information that could lead to discrimination. For example, in AI systems, it is no longer allowed to include skin color or gender in decisions make their judgment in that regard neutral. However, it's quite challenging to discover all the information in the data that could lead to a biased result. For example, an address can reveal whether one is more likely to be poor or wealthy if poorer or wealthier people live in a particular neighborhood. In addition, structured additional data (so-called metadata) in datasets can also contain information that allows conclusions to be drawn about individuals, without the user noticing. Therefore, it is essential to thoroughly examine which data the system should learn from when creating AI systems. Also, the results of the AI system must be critically reviewed regularly to ensure they are neutral. Therefore, attempts are made to develop AI systems in a way that all decisions are comprehensible. Besides, it should be possible to explain on which information the decisions are based. This leads to AI systems that are transparent and fair. In addition, society needs rules and laws for what tasks AI systems can be used and what decisions they can make to protect people. These are currently being developed by politicians in Germany and Europe.



And what does an AI system say about the question, „Are AI systems racist?“ ChatGPT responds as follows:

„Artificial intelligence (AI) systems themselves are not racist because they have no emotions or prejudices. However, AI systems can produce racist results due to biases in the data with which they were trained. When AI systems are trained on data that reflects unconscious or even conscious biases as they exist in society, they can unintentionally make racist or discriminatory decisions. For example, AI systems trained on job application data may automatically prefer male candidates if they have previously selected more male applicants because this is the basis of their training. It is therefore important to pay attention to careful data selection and cleansing in the development of AI systems to minimize biases and discrimination. In addition, appropriate monitoring and controlling of AI systems should be carried out to ensure that they remain fair and unbiased.“





SOURCES

Graphic "Rejected due to: Discrimination!":

created by: Michaela Müller-Unterweger, Photos: <https://pxhere.com/de/photo/706486>, <https://pxhere.com/de/photo/706040>, beide CCO-Lizenz, <https://pixabay.com/vectors/rubber-stamp-rejected-denied-vector-2054006/>, Fonts: Calibri, Really Free Font, CC-BY-ND Ermin Dwi Alita, Annisa Aulia Rasyidah, and Herbanu Tri Sasongko

Racist Google algorithm

<https://www.zeit.de/digital/internet/2015-07/google-fotos-algorithmus-rassismus>

ChatGPT

created using <https://chat.openai.com/>, April 2023

SOURCES (UN)FAIRE AI?! - APP

Icons

Cuffs: https://www.flaticon.com/free-icon/handcuff_4435702?term=handcuff&page=1&position=2&origin=search&related_id=4435702

Bag: https://www.flaticon.com/free-icon/portfolio_1063376?term=job&page=1&position=1&origin=search&related_id=1063376

Building: https://www.flaticon.com/free-icon/buildings_558411?term=flat&page=1&position=3&origin=search&related_id=558411

School: https://www.flaticon.com/free-icon/classroom_3197831?term=school&page=1&position=20&origin=search&related_id=3197831

Data Set: https://www.flaticon.com/free-icon/dataset_8003249?term=dataset&page=1&position=3&origin=search&related_id=8003249

Home Button: https://www.flaticon.com/free-icon/home_25694?term=home&page=1&position=1&origin=search&related_id=25694

Portraits Testpersons

Alan: <https://pixabay.com/de/photos/schwarzer-mann-afroamerikaner-3532117/>, Photo by jidaley

Maria: <https://pixabay.com/de/photos/frau-frauen-b%20cro-arbeit-gesch%20a4ft-2773007/>, Photo by ernestoeslava

Carmen: <https://pixabay.com/de/photos/frau-iranische-frau-alte-frau-iran-7262808/>, Photo by Javad_esmaeili

Fabian: <https://pixabay.com/de/photos/mann-rauchen-rauch-zigarette-4132758/>, Photo by 10634669



SOURCES **(UN)FAIRE AI?! - APP**

Protraits Testpersons

Stefanie: <https://pixabay.com/de/photos/frau-brille-business-frau-fachmann-1254454/>,

Photo by ukieiri

Almuth: <https://pixabay.com/de/photos/arbeiten-gesch%a4ftsfrau-weiblich-1219889/>, Photo by helpsg

lich-1219889/, Photo by helpsg

Ramona: <https://pixabay.com/de/photos/marketing-gesch%a4ftsmann-3468598/>,

Photo by madartzgraphics

Artur: <https://pixabay.com/de/photos/m%a4nnliches-portr%a4t-mann-offiziell-1156365/>, Photo by akintans

ziell-1156365/, Photo by akintans

Jürgen: <https://pixabay.com/de/photos/mensch-person-mann-nachbar-3429797/>,

Photo by 652234

Adam: <https://pixabay.com/de/photos/%a4lterer-herr-hunde-h%bcndchen-mann-1281284/>, Photo by brenkee

mann-1281284/, Photo by brenkee

Familie Stretz: <https://pixabay.com/de/photos/familie-hund-portr%a4tk%bcche-liebe-7071153/>, Photo by Michelle_Raponi

k%bcche-liebe-7071153/, Photo by Michelle_Raponi

Familie Nassar: <https://pixabay.com/de/photos/familie-spielen-spa%9f-1%a4chelt-kind-2901211/>, Photo by Chepearroyo

1%a4chelt-kind-2901211/, Photo by Chepearroyo

Paul: <https://pixabay.com/photos/boy-teenager-male-young-portrait-4433619/>,

Photo by djedj

Malaika: <https://pixabay.com/de/photos/m%a4dchen-kind-schwarze-afrikanisch-486950/>, Photo by Counselling

nisch-486950/, Photo by Counselling

Melissa: <https://pixabay.com/de/photos/m%a4dchen-mensch-h%bcbsch-weiblich-707813/>, Photo by Axe77

lich-707813/, Photo by Axe77

Nikolaj: <https://pixabay.com/photos/homework-school-child-homeschooling-2803933/>,

Photo by anaterate

