



AI'S BIAS CHALLENGE

How stereotypes are shaping Artificial Intelligence and what to do about it

By IMD Professor Bettina Büchel

IMD
Chemin de Bellerive 23
PO Box 915,
CH-1001 Lausanne
Switzerland

Tel: +41 21 618 01 11
Fax: +41 21 618 07 07
info@imd.org
www.imd.org

Headlines about bias in Artificial Intelligence keep on coming.

This week the AI Now Institute, based at New York University, [released a report calling the field of AI a “diversity disaster”](#).

Last week [Facebook was in the hot seat](#) for allegations of bias in its ad-targeting algorithms.

That’s on top of a recent privacy-related [law suit Facebook is up against in Germany](#).

We’ve [known for some time](#) that with Artificial Intelligence there is a very real risk of reinforcing society’s gender equality problems.

How AI becomes biased

The way machines learn, a subfield of AI, involves feeding the computer sets of data –whether it’s in the form of text, images or voice – and adding a classifier to this data. If search engines are being used to train the machine learning systems, then the images used to portray what jobs people hold or what activities they do are often based on societal stereotypes. An example would be showing the computer an image of a man as a board member or a STEM scientist, or a woman as a nurse or teacher.

Over time and with many images, the computer will have learned to recognize similar images and classify them according to past patterns. One example of [AI already in use](#) that shows the potential for bias is how it is used to assess credit risk before issuing credit cards or small loans. An AI algorithm could easily be used to rule out a group perceived to be from a less financially stable demographic, like single women, for example. A person could also be rejected for a mortgage if they currently live in a low-income area.

Another potential trouble area is job candidate selection. LinkedIn, for instance, had an issue where highly-paid jobs were not displayed as frequently for searches by women [as they were for men](#) because of the way its algorithms were written. The initial users of the site’s job search function were predominantly male for these high-paying jobs, so it just ended up proposing these jobs to men.

My colleague Maude Lavanchy [has written about how Amazon abandoned its experimental artificial intelligence](#) recruiting tool after discovering it discriminated against women.

AI algorithms are trained to observe patterns in large data sets to help predict outcomes. In Amazon’s case, its algorithm used all CVs submitted to the company over a ten-year period to learn how to spot the best candidates. Given the low proportion of women working in the company, as in [most technology companies](#), the algorithm quickly spotted male dominance and thought it was a factor for success.

Because the algorithm used the results of its own predictions to improve its accuracy, it got stuck in a pattern of sexism against female candidates, Maude wrote.

Don’t be fooled by Artificial Intelligence

The origin of these biases stem from sample bias, stereotypes, measurement distortion or the algorithmic model itself (although this one is not as likely). Sample bias comes from giving the system data that does not fully reflect the actual choice set. Board positions within companies are primarily men so if the system is given only examples of men to learn from, then the existing bias will be recreated in the future since the data being used comes from today's reality.

What to do in the meantime

Until we can fix AI's bias problem humans have an important role to play in designing new AI systems to avoid bias and prevent discrimination. We, as human decision-makers still need to be wary of AI and must have the final say on whether certain predictions get acted upon, if we want to prevent discrimination in many important areas of business.

*[Bettina Büchel](#) is Professor of Strategy and Organization at IMD business school. Her forthcoming book is titled *Strategic Agility: The art of piloting initiatives*. She is the Director of TransformTECH, a program which helps executives use cutting-edge technology to explore new business opportunities.*