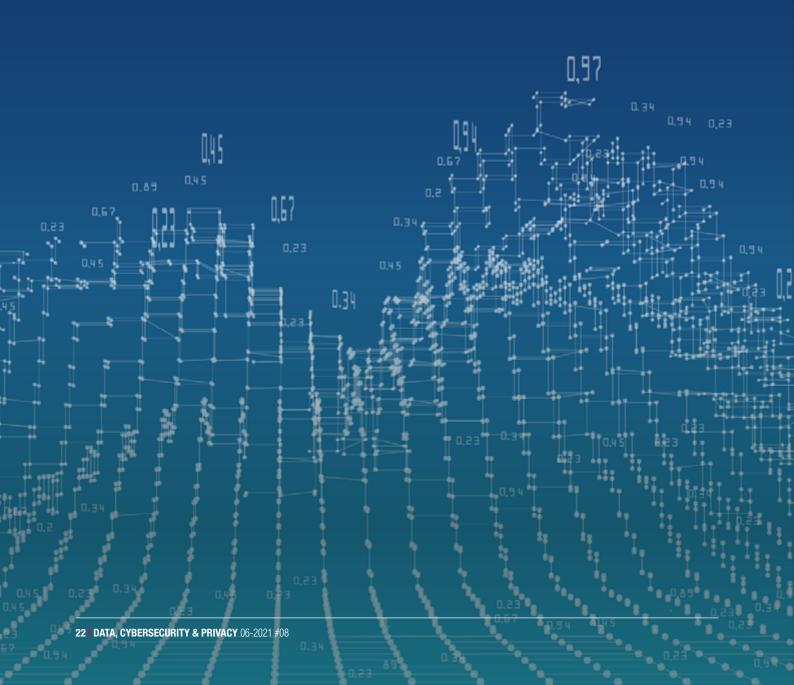Quirine Eijkman - Netherlands Institute for Human Rights

# Algorithms increase *ánd* decrease a fair chance to get a job

**Everyone wants a fair chance to get a job. We want employers and intermediaries to consider our talents, skills and experience and not, for example, our age, ethnicity, gender or sexual preference. More and more employers are using algorithms to be able to select quickly. Research from the Netherlands Institute for Human Rights suggests that the use of algorithms can increase, but may also reduce the chance of discrimination.**

### Increased use of algorithms

'Unfortunately, you have not achieved the required minimum score, which means that we will not continue the procedure. Thank you very much for your application and enthusiasm for this position.'

Perhaps you have received a message like this after applying for a job: a rejection prepared by an algorithm. Increasingly, job applicants are faced with automated selection procedures in which employers make use of algorithms. Algorithms can unfortunately have a discriminatory effect. This kind of discrimination is often not very visible and is hidden behind computer code. After all, how can you find out if you have been rejected by a discriminatory algorithm when you receive a message like the one above? In its supervisory role as the Netherlands' national human rights institution the Institute for Human Rights monitors developments concerning the use of algorithms in recruitment and selection. Last year, the Institute presented a number of findings in a research report. By using examples, the Institute explains the situation regarding discrimination and algorithms. We map out the causes, risks, but also the potential of algorithms to prevent discrimination.

### Bias in the algorithm

In this literature based study two causes of algorithmic discrimination are assessed: bias in the algorithm and data bias. Bias in the algorithm may be introduced when choosing variables for the algorithm or while designing the algorithm. Variables that at first glance do not appear to be discriminatory at all, may still lead to discrimination. A variable like 'years of uninterrupted employment', for example, can be an indication of good performance, but also an indication of gender, because women are more likely to have interruptions in their careers. Even without explicitly providing information about someone's

> ## "Algorithms are sometimes simply too complex to be transparent."

gender, race, religion, or other protected attributes, an algorithm can discriminate certain groups.

### Bias in data

Data bias occurs because data can reflect prejudice that exists in society. If biased data is subsequently used to train or edit an algorithm, it can lead, in turn, to bias in the algorithm. This is what happened for instance with a selection algorithm used by a UK university that preferred men without a migration background. This algorithm was trained based on data from a period when the university admitted few women and migrants.

Another example concerns LinkedIn. In 2016, LinkedIn was the centre of some controversy. It turned out that when users searched for a female contact, LinkedIn's search engine assumed they were searching for a man, suggesting a male name that looked like the female name they entered. A search for 'Stephanie Williams', for example, would return 'did you mean Stephen Williams?'.
The same happened with other female names: Danielle was assumed to mean Daniel and Alexa became Alex. According to LinkedIn, these suggestions were given based on users' search data; the names were simply searched more often. LinkedIn has meanwhile modified its search feature.

### Algorithm acquires bias

Moreover, an algorithm that is already in use may also acquire bias based on user data. One algorithm, for example, pushed online adverts for engineering vacancies mainly to men because past adverts received more clicks from men than women. Hence, women had fewer chances to click these job adverts, while there may very well have been women who would be suited for and interested in these jobs.

Bias may also arise because certain groups are underrepresented in the data. People with certain impairments and elderly people, for example, are less likely to use social media, meaning that their preferences are underrepresented in the training data for the algorithm.

### Heightened risk of discrimination

Although people are not necessarily good at leaving their biases and prejudice out of hiring decisions, there are

four risks that can further exacerbate discrimination due to the very nature of algorithms.

### 1. Black box

It can often be unclear and intransparent how algorithms arrive at decisions. Algorithms are sometimes simply too complex to be transparent. This complexity makes it impossible to trace what happened, even for those who designed the algorithm. The main reason for that is that humans are simply no longer able to interpret the code. On top of that, companies tend to be reluctant to disclose their algorithm because they want to protect their commercial interests and business secrets.

This opacity of algorithms, which is also referred to as the black-box effect, means that job applicants are often unable to find out why they were rejected for a job in the first place. This makes it hard or even impossible for applicants to prove they were discriminated against. After all, how can you know that you have been rejected by a discriminating algorithm if you have only received a rejection e-mail without further motivation. Moreover, the black box makes it easier to conceal (intentional) discrimination.

### 2. Automation bias

People generally associate computers and algorithms with rational decision making and faultlessness, which is referred to as automation bias (by people). Automation bias leads to the risk of recruiters having an unjustifiably high level of confidence in algorithms without being alert to discrimination risks. Precisely due to the time pressure in the recruitment and selection process, there is a great risk of automation bias.

### 3. Systematising discrimination

Thirdly, algorithms can further embed and systematize discrimination in society. Algorithms are, in principle, consistent. When an algorithm makes decisions based on predefined parameters that contain a bias, it will, in principle, apply these parameters to all decisions. People, on the other hand, while in no way free of prejudices, have freedom of choice and intuition that enable them to become aware of their biases and to unlearn them.

> "Employers need to be extra careful when using this technology."

### 4. More options

Fourthly, algorithms combined with new data collection technology heighten the risk of unequal treatment because they are able to detect highly complex correlations. This leads to a greater risk of groups being classified based on all kinds of neutral data that would be classified under the traditional grounds of discrimination, such as whether someone is rich or poor.

Algorithms also have the capability to discover more correlations between seemingly neutral information and grounds of discrimination, as a result of which they can contain an unidentified bias against certain groups. In short, there are more ways to differentiate between people and, consequently, to discriminate.

### Using algorithms to fight discrimination

Despite the risks for recruitment and selection, the Institute also sees the potential of algorithms to be more objective than people and to be used as a tool in the fight against discrimination. However, technology that removes bias from algorithms is still in its infancy and involves a lot of risks. Further research into ways to improve algorithms and to include equal treatment considerations in the design of an algorithm, i.e. non-discrimination by design, is essential to prevent algorithmic discrimination.

Another challenge is that bias-free technology will need to have access to information about gender, race, or other grounds of discrimination to correct itself. The question is whether this is desirable, given the risk of direct discrimination and because this kind of data can, under the current General Data Protection Regulation (GDPR), only be used if stringent requirements are met.
On top of that, it is difficult to draw a line between correcting bias in an algorithm and giving preferential treatment, which is also subject to strict legal rules.

### Awareness and research

It is essential that algorithm designers and Human Resource (HR) professionals are aware of the risk of discrimination when using algorithms for recruitment and selection. The use of algorithms in the recruitment and selection process creates new challenges in terms of human rights. This study suggests that the use of algorithms in recruitment and employment selection processes come not only with challenges for the protection of the right to privacy and data protection, but also involve discrimination risks.

> "Job applicants are urged to report any case where they believe they have been discriminated against by a recruitment and selection algorithm."

The question remains, however, how the requirements under (EU) anti-discrimination law apply to algorithms. This was also noted by the researchers behind a recent study by the Netherlands Organisation for Applied Scientific Research (TNO) entitled Digital labour market discrimination (Digitale Arbeidsmarktdiscriminatie), which was assigned by the Dutch Social Affairs and Employment Inspectorate (Inspectie Sociale Zaken en Werkgelegenheid). The researchers noted that there is little case law and research on the subject. The Institute intends to assess this question in a future follow-up study.

### Transparency from designers
Based on this study, the Institute makes several recommendations. We recommend developers to inform clients about the potential risks involved with using algorithms for recruitment and selection. Also, they need to disclose the variables based on which applicants are assessed and do so in an accessible and comprehensible manner. Furthermore, we advise designers to regularly perform validations to check whether their software does not lead to discrimination.

### Information duty for employers
Employers need to inform job applicants that an algorithm will be used in the selection process in time and explain in layman's terms the function of the algorithm in the process. Besides this they need to make sure they do not use recruitment technology when it is unclear which criteria are used to judge applicants.

### Job applicants report discrimination
Job applicants are urged to report any case where they believe they have been discriminated against by a recruitment and selection algorithm, even if it is merely a suspicion on their part. The local discrimination reporting centres can play an important role in this regard. Naturally, everyone who is discriminated in

the Netherlands can contact the Institute to request a ruling on a potential case of discrimination, to report discrimination, or to ask any question they may have about discrimination.

### Programme Digitalisation & Human Rights
The Institute executed this research in the context of their strategic programme Digitalisation & Human Rights. Digitalisation can threaten human rights. This not only concerns privacy and data protection, but also other human rights such as the right to equal treatment and a fair trial. The Institute contributes to the continuation of digitalisation without people being disadvantaged, in order to make the Netherlands a society in which everyone can participate.

**In short:**
- Algorithms can discriminate due to bias in the algorithm itself or in the data. The use of algorithms can worsen the effects of discrimination.
- This is partly due to the fact that algorithms are a black box, making discrimination harder to detect.
- Algorithm designers and employers must be transparent about how algorithms work.
- The government must provide more information to job seekers, employers and algorithm designers.



**About the author**
Quirine Eijkman is deputy president research & advice at the Institute for Human Rights and Chair of the Research group Access to Justice at the Centre of Social Innovation (KSI) of the HU University of Applied Sciences Utrecht.