DRITTER GLEICHSTELLUNGSBERICHT

Algorithms and Discrimination

Fact Sheet 5

Digitalised discrimination

"The [technical] instruments we design derive their values from the society in which they are embedded. ... In a rational society, the computer would be nothing but good." (Joseph Weizenbaum)

Projections calculated by computers or algorithmic systems are now often used in areas where crucial life decisions are made: in staff selection, in granting loans, in decisions about social and health benefits, or even in probation or prison sentences.

Algorithmic systems consist of dozens to thousands of complexly interlinked mathematical procedures, called **algorithms**. These are often designed by means of evaluating huge datasets. In the process, parameters of the mathematical prediction models are created in the form of equations from statistical data. Algorithms are written in programming languages and referred to as **software**. This software can be distributed over several spatially separated, networked computer and storage systems.

» The start-up company Retorio uses software to create, for instance, personality profiles of job applicants based on their facial expressions and gestures in video-based job interviews. The software is supposed to be less biased than a human interviewer. However, a test by German public broadcaster Bayerischer Rundfunk showed that this personality model changes when a candidate switches hairstyle, glasses or clothing, sits in front of a different background, or adapts brightness values. In addition, the values for the personality trait "conscientiousness" changed considerably when the same test person wore a headscarf but did not change anything else about her personal behaviour.

» In the United States, an algorithmic system is used to decide on more intensive care for patients. This system initially calculated the need for treatment based on the costs that patients had previously caused within the health system: the higher the costs, the higher the need for treatment. Millions of Black people were disadvantaged by this algorithm. The reason: the costs they incurred are lower, as they often have less access to the health system and thus make less use of preventive measures, for instance.

Algorithmic systems may discriminate against people on the basis of their gender, sexual identity, ethnic origin, skin colour, disabilities, level of education, social status, or religious and ideological beliefs.

IN THIS FACT SHEET

Garbage in – Garbage out
Algorithmic systems that provide forecasts or decision-making support often require large amounts of statistical data, called 'training data'. If the selection of training data is not representative of all those persons subsequently affected, if it reproduces stereotypes, or if it is based on false causal assumptions (example 2, page 1), a **bias effect** occurs. With such data, reality is thus presented in a biased and prejudiced way.

How does discrimination by algorithms emerge?

Discrimination by algorithmic systems is often not intended by the developers. However, even supposedly objective procedures may entail considerable risks of discrimination. Discrimination by algorithms can arise in particular

» if too little research is done on the future area of an algorithmic system's application. This includes the precise observation of the working methods of future users, needs assessments, etc.,
» if biased data are used for modelling the system,
» if prototypes or intermediate versions of the software are not tested or are tested on the basis of too few criteria and only by people involved in the development – while future affected persons are not included,
» if development teams lack the necessary sensitivity and expertise to recognise discrimination risks,
» if experts on discrimination risks – e.g. gender equality officers or experts on accessibility, gender and diversity – are not involved in the development,
» and if the objective of a task to be designed is fundamentally problematic, e.g. issuing further training measures at the employment office on the basis of the chance of re-entry into the workforce – calculated on the basis of age, gender, place of residence, care obligations and nationality, as in the algorithm used by the Austrian labour market service.

How can discrimination by algorithms be prevented? Existing regulations and necessary adjustments

The Expert Commission for the Third Gender Equality Report of the German Federal Government has developed recommendations geared towards reducing discrimination by algorithmic systems for policymakers and businesses. These recommendations focus in particular on the use of algorithms in staff selection, as discrimination in such decisions significantly impairs people's capabilities in working life. The German Basic Law, the EU General Data Protection Regulation (GDPR), and the German General Act on Equal Treatment (AGG) already protect consumers from discrimination in work contexts. However, amendments are necessary to increase protection against discrimination, especially discrimination caused by algorithmic systems:

Processing of gender-related data

The GDPR prohibits the processing of categories of data that are particularly susceptible to discrimination and that reveal "racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership" (article 9 (1) GDPR). What should be amended:

Gender-related data may only be processed in exceptional cases

» Data revealing gender or other legally protected characteristics such as sexual orientation must be specially protected – just as the specific categories of data within the meaning of article 9 GDPR. The processing of this data must be prohibited as a matter of principle and may only be permitted in exceptional cases which are themselves regulated by law.

Mandatory auditing

The AGG prohibits discrimination in working life "on grounds of race or ethnic origin, gender, religion or belief, disability, age or sexual identity" (articles 1, 2 (1, 7) AGG). To enforce this, algorithmic systems must be thoroughly audited and checked. This can be done under the provisions of the GDPR. Already today, the body responsible for data processing must assess the consequences for the protection of personal data in advance of a particularly risky data processing operation. However, more regulation is needed:

Obligation to carry out data protection impact assessments with regard to discrimination

» Discrimination risks resulting from the use of algorithmic systems must be explicitly monitored within the framework of data protection impact assessments.

» When using algorithm-based decision-making procedures, a data protection officer must be appointed within the company. If available, gender equality officers must be involved in the development and implementation of algorithm-based procedures in the company.

Including audit obligations for the use of algorithmic systems in the AGG and GDPR

» Employers must ensure that algorithmic systems for staffing decisions are designed in a way that disadvantages according to article 1 AGG are ruled out.
» Employees and applicants must be granted the legal right to demand an immediate review of the algorithmic decision-making process by company data protection officers.
» Algorithmic systems must be reviewed at least once a year. The review must be carried out in commissions with equal representation of the (staff) management and employees or their representatives.

For internal audits, **audit guidelines** can be used, for example the audit guide by Algorithmwatch (2020) „Automatisierte Entscheidungen und Künstliche Intelligenz im Personalmanagement. Ein Leitfaden zur Überprüfung essenzieller Eigenschaften KI-basierter Systeme für Betriebsräte und andere Personalvertretungen" ("Automated decisions and Artificial Intelligence in human resource management. A guide to reviewing essential features of AI-based systems for works councils and other staff representatives') at https://algorithmwatch.org/de/auto-hr/leitfaden/.

Training the works council and staff representatives

» To comprehend violations in the use of complex algorithmic systems, actors of collective interest representation must have digitalisation-related competences as well as discrimination sensitivity. It must therefore be ensured that staff representatives and works council members attend appropriate training.

Risk assessment and grading

At the end of 2019, the German Federal Government's data ethics commission proposed a "criticality pyramid" with five risk levels aimed at assessing the potential harm of algorithmic systems for individuals and for society. The upcoming EU regulation "Artificial Intelligence Act" does take up a risk-adapted approach. However, the assessment of discrimination risks must become more concrete and binding in this case.

Risk assessment for software systems

» The recommendations of the data ethics commission should be incorporated into binding regulations.
» The Expert Commission does explicitly not exclude the classification of certain algorithmic systems in the area of staff selection procedures in the highest criticality level. Their prohibition may thus possibly be advisable.

Duty of information

According to the GDPR, data subjects must be informed about the use of an algorithmic system, for example with regard to staff selection. This means:

Clarifying and specifying the right to information under data protection law

» The existing data protection rights to information (in particular article 15 GDPR) must be specified as to the use of algorithmic systems in staffing decisions. Data subjects must be comprehensively informed about the use of algorithmic systems and about how these work.

Non-transparency

The exact modes of operation and data used by the technologies existing on the market are usually not transparent and/or traceable. For example, an applicant cannot easily find out how her/his online behaviour is evaluated on job portals like LinkedIn. What influence does this behaviour have, for example, on recommendations to potential employers? Such information is often treated as corporate secrets. This must change:

Safeguarding transparency of algorithmic systems

» Companies that develop algorithmic systems must ensure that their technical specifications, programming guidelines, requirements catalogues, documentation, and source codes are disclosed.

Victimless discrimination

Discrimination often goes unnoticed by those affected, for example when female users never find out that they were not shown the new advertisement for a job as a lorry driver. In social networks such as Facebook, the targeting of job advertisements on the basis of gender, called gender targeting, has repeatedly been proven. Potential countermeasures include:

Introducing a right for associations to sue

» Anti-discrimination bodies and associations should be equipped with further competences and powers, specifically with the right for associations to take legal action against discrimination without any identifiable victim(s).
» Such a right would relieve employees and job applicants from having to take personal action against their employers in the course of their employment.
» In addition, there should be a lump-sum compensation for damages that can be claimed by anti-discrimination bodies.

Generalisations

Automated decision-making by means of algorithmic systems turns a blind eye to individual circumstances and single cases. How, for example, are "gaps in the CV" recorded and individually interpreted? Who takes responsibility for unjustified or incomprehensible machine assessments?
The GDPR does provide for a specific right to object to the processing of personal data (article 21) and prohibits exclusively automated decisions. However, there are wide-ranging exceptions, including when data subjects consent or an automated process is in fact necessary for the purposes of fulfilling a contract. A next step is therefore:

Improving employee data protection

» In addition to article 22 of the GDPR, a national employee data protection law should furthermore prohibit the complete automation of parts of a decision-making process without taking individual circumstances into account.
» In addition, a reorganisation of employee data protection should also include a requirement that this prohibition cannot be waived through the consent of employees or job applicants.

Further reading

» Chapter B.III.3 in the Expert Opinion part of the Third Gender Equality Report of the German Federal Government, available (in German) at: https://www.bmfsfj.de/gleichstellungsbericht
» Agency for the Third Gender Equality Report of the German Federal Government (2021): Shaping digitalisation in a gender-equitable way. Summary of the Expert Opinion of the Third Gender Equality Report of the Federal Government. Berlin: Agency for the Third Gender Equality Report. Download at: https://www.dritter-gleichstellungsbericht.de/de/topic/50.english.html
» Jaume-Palasí, Lorena/Kloiber, Julia/Lindinger, Elisa (2020): AI Powered Recruiting? Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst. Expert opinion for the Third Gender Equality Report of the German Federal Government, available (in German) at: https://www.dritter-gleichstellungsbericht.de/de/topic/62.expertisen.html
» Orwat, Carsten (2019): Diskriminierungsrisiken durch Verwendung von Algorithmen, available (in German) at: https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Expertisen/studie_diskriminierungsrisiken_durch_verwendung_von_algorithmen.html

IMPRINT:

Fact Sheet by the Agency for the Third Gender Equality Report
Publisher: Institut für Sozialarbeit und Sozialpädagogik e.V.
Agency for the Third Gender Equality Report of the German Federal Government
Sebastian Scheele and Dr. Ulrike Spangenberg (heads of management)
Lahnstraße 19, 12055 Berlin
www.dritter-gleichstellungsbericht.de
Currently effective version of: October 2021
Year of publication: 2022

ISS Gemeinnütziger e. V.

Funded by:
Federal Ministry for Family Affairs, Senior Citizens, Women and Youth