Federal
Anti-Discrimination
Agency

The Federal Anti-Discrimination Agency's research at a glance:
Risks of Discrimination through the Use of Algorithms

Overview of the study

The current study examines how the use of algorithms can lead to unjustified discrimination when they are used to differentiate between individuals. Among other things, algorithms are used to automate processes. That means that either recommendations for decisions are automatically derived based on computer-assisted data processing and analysis or decision-making rules are executed completely automatically. Through automated differentiation of individuals on the basis of legally protected characteristics or with the help of surrogate data, the scope for human interpretation and action is curtailed. At the same time, computer-generated stereotypes and the potential for discrimination are created. By means of example cases, the study analyses how risks of discrimination arise, which impacts they may have on society and it derives considerations for the need of action and possible measures to avoid discrimination.

Authors, year of publication and title of the study
Institute for Technology Assessment and Systems Analysis (ITAS), Karlsruhe Institute of Technology (KIT). Orwat, Carsten (2019): Risks of Discrimination through the Use of Algorithms.

Results

Reasons for the risks of Discrimination trough the Use of Algorithms.

- In the **development of algorithms and models**:
    - ➢ Risks arise from the labelling of categories for the classification of individuals especially where those rely on subjective assumptions and interpretations. (example: characteristic: "suitability for a company")
- Through **biased (training) data sets:**
    - ➢ If data sets are incomplete, no longer up-to-date or stem from situations in which there was or is unequal distribution or discrimination against individuals, this can lead to certain groups being over- or under-represented.
    - ➢ Due to the substitution of legally protected characteristics with seemingly "neutral" variables, risks of indirect discrimination can arise if there is a correlation between the variables and the legally protected characteristics. (example: correlation between place of residence and ethnicity)
    - ➢ Data mining and machine learning processes use more variables compared to "traditional" statistical methods. This increases the risk of (undetected) correlations.

- **In online platforms:**
  - ➢ Here, algorithms help users rate and select one another. This can lead to restricted access to certain interactions and transactions of certain users.
  - ➢ In cases where algorithms are based on evaluations and rankings of other users, social inequalities can multiply.
  - ➢ Pricing and market mechanisms (e.g. auction mechanisms) which are used in advertising placements and customer selection might also be responsible for risks of discrimination.
- The use of computer-based systems **can conceal cases of deliberate discrimination.**

Social risks of algorithm-based differentiation
- Statistical discrimination does not categorise individuals based on their actual characteristics. Instead categorisation results from the processing of group data. Thus, the use of stereotypes generated by data processing determine the outcomes of decision-making processes. Unfair generalisation can occur, especially in "atypical" cases.
- When they are being categorised and assigned to a certain group based on algorithms, individuals affected do not have the chance to agree or to disagree with the treatment they are being subjected to. This also poses a threat to the freedom of personal development, the right to self-expression and the protection of human dignity.
- The risks connected with economic-rational differentiation might add up to cumulative disadvantages because algorithmic assessments and actions based on them could concentrate on individuals or groups who are already disadvantaged and thus appear more frequently in the data collection.
- Equality and social political objectives may be weakened if societal considerations on differentiation (made possible with the help of algorithms and seemingly economically viable) proceed in a one-sided way in favour of efficiency and at the expense of equality.

Need for action and possible measures
- **Need for action in society:**
  - ➢ There need to be processes for societal consideration and balancing which take into account what is being gained in terms of differentiation and efficiency but take into account above all their social distribution. They should also lead to determining which applications of algorithm-based differentations are acceptable for society.
- **Legal Adjustments:**
  - ➢ The data protection legislation should be corrected and specified in view of concrete information requirements and with regard to the concept of informed consent. Thus, making the intended effects and risks of discrimination linked to the use of algorithms foreseeable. The regulation of automated decision systems also needs to be clearly defined.
  - ➢ The Equal Treatment Act regulations concerning the burden of proof need to be addressed as it is difficult for persons concerned to understand and retrace the disadvantages caused by algorithms. Where appropriate, it should be mandatory to document the elements and results of algorithm-based procedures.
  - ➢ Moreover, a collective legal protection by means of a right to collective action should be established. It should be examined whether the Equal Treatment Act and the catalogue of protected characteristics need to be expanded.

- **Possible courses of actions for the Federal Anti-Discrimination Agency (FADA):**
  - ➢ In order to prevent discrimination, the developers and users of algorithms should be advised on the risks of discrimination arising from these.
  - ➢ It should be made mandatory that the FADA be involved in the decision-making process when public bodies source algorithms.
  - ➢ Opportunities for the direct analysis and the testing of algorithms and software systems should be created with the help of relevant IT skills.
  - ➢ The FADA should point out situations, groups or characteristics vulnerable to discrimination and should have a say in the interpretation and the assessment of the lawfulness of differentiations.

Additional information

You can access the study "Risks of Discrimination through the Use of Algorithms" => here.

Contact: Federal Anti-Discrimination Agency
                Glinkastraße 24
                10117 Berlin, Germany
                Phone: +49 (0) 3018 555 – 1855
                Initial legal advice – e-mail: beratung@ads.bund.de
                General enquiries – e-mail: poststelle@ads.bund.de