D. Demmler, D. Krupka, H. Federrath. (Hrsg.): INFORMATIK 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 

# ‘Not all algorithms!' Lessons from the Private Sector on Mitigating Gender Discrimination

Mareike Winkler [1], Sonja Köhne[2], Miriam Klöpper[3]

**Abstract:** In the public sector, the use of algorithmic decision-making (ADM) systems can be directly linked to crucial state assistance, such as welfare benefits. Prominent examples such as an algorithm of the Public Employment Service Austria, that predicted below-average placement chances for women, underline the high risks of systematic gender discrimination. The use of ADM is rather novel in the public sector. The private sector, on the other hand, can resort to a relative wealth of experience in adopting such algorithms and dealing with algorithmic gender discrimination, for example in recruiting. Based on empirical examples our paper 1) explores how gender is currently considered in the development of ADM for the public sector, 2) highlights the potential risks of algorithmic gender discrimination, and 3) analyzes how the public sector can learn from the experience of the private sector in mitigating these risks.

**Keywords:** eGov; algorithmic decision-making; automation bias; algorithmic bias; gender discrimination

## 1 Introduction

Advocates and vendors of algorithmic decision-making (ADM) systems often claim that the use of these systems potentially results in more rational, objective and fair decisions [Tu18]. Recent examples from the private and public sector have proven otherwise, as the deployed algorithms inherit human biases. One widely discussed example is a hiring algorithm of technology company Amazon that was trained with historical data of successful hires of the previous ten years. As the training data reflected a male-dominated IT industry, the algorithm inappropriately predicted that male candidates would be a better fit than their female competitors, thus systematically disadvantaging women [Da18]. According to the company itself, this algorithm was never used in practice but has since been widely cited in the media discourse around algorithmic bias.

[1] Carl von Ossietzky University of Oldenburg, Institute for Social Sciences, 26111 Oldenburg; Denkfabrik Digitale Arbeitsgesellschaft at the German Federal Ministry for Labour and Social Affairs*, Wilhelmstraße 49, 10117 Berlin, mareike.winkler@uni-oldenburg.de, https://orcid.org/0000-0002-8608-9936

[2] Alexander von Humboldt Institute for Internet & Society, research group Innovation, Entrepreneurship & Society, Französische Straße 9, 10117 Berlin; Denkfabrik Digitale Arbeitsgesellschaft at the German Federal Ministry for Labour and Social Affairs*, Wilhelmstraße 49, 10117 Berlin, sonja.koehne@hiig.de, https://orcid.org/0000-0002-2793-0529

[3] FZI Forschungszentrum Informatik, department Innovation, Strategy and Transfer, Haid-und-Neu Straße 10-14, 76131 Karlsruhe, kloepper@fzi.de, https://www.fzi.de/team/miriam-klopper/

*The views expressed in this chapter are those of the authors.

doi:10.18420/inf2022_110

Another prominent example is an ADM system which was used in the Public Employment Service (PES) Austria, an agency that is responsible for reintegrating unemployed into the labor market. To cope with high workloads and to allocate resources more efficiently (e.g. vocational training measures), the PES Austria introduced an ADM system that calculated the chances for the unemployed to find a job within a certain period of time [HW18; Or20]. Including the variable 'gender' negatively affected the calculated job chances of unemployed women, while it had no effect for unemployed men. This, in turn, affected how resources for employment-supporting measures were allocated. The threats of ADM systems to gender inequality are underlined by many more international examples ('not all algorithms but still too many'[4]). Some experts claim that the risks might even outweigh the potential benefits [WH22; Ad22]. Thus, the deployment of algorithms to decide on the distribution of goods and services would pose a challenge to the general principles of distributive justice.

Against this background, the research question we seek to answer in this paper is: How can risks of algorithmic gender discrimination be mitigated in the public sector? Therefore, the paper proceeds as follows. It 1) explores how gender is currently considered in the development of ADM for the public sector, 2) highlights the potential risks of gender discrimination, and 3) analyzes if and how the public sector can learn from the experience of the private sector.

## 2 ADM and gender discrimination

### 2.1 Potentials of ADM in the public sector

The public sector in Europe is increasingly using ADM systems in a variety of application contexts [Al20], such as for predicting recidivism in criminal justice [Ge22] or identifying social welfare fraud [El21] (for a catalogue of ADM systems in Germany see [AW19]). ADM refers to an algorithmic system that supports or substitutes human decision-making processes [Kr20]. It stands for a variety of algorithmic systems that are developed for a single specific task, e.g., the above-mentioned identification of social security fraud. An ADM system produces values that represent a certain category (i.e., a classification or ranking) or a risk or probability (e.g., the likelihood that a person will obtain welfare benefits by false pretenses). To do this, the algorithmic system (more precisely, the model) processes a given data input (e.g., income, household size, etc. of the welfare recipient) and predicts the probability of a given output (e.g., welfare fraud). The algorithm is either given – in other words, certain features in the data lead to the conclusion of welfare fraud – or it derives certain features based on historical data (e.g.,

[4] The phrase is an adaptation of the Internet meme 'Not all men are like that'. The meme satirically illustrates how some men deflect discussions about misogyny, etc., insisting that they personally are not to blame. (See also: time.com)

on eligible and ineligible welfare recipients of the last 20 years). The algorithmic component may be based on expert systems, machine learning (ML) or other data-intensive technologies [Le21]. Levi and colleagues (2021) point out that "machine learning, impressive as it may sound, describes what are fundamentally statistical techniques for fitting (sometimes very complicated) models to data". Nevertheless, it should also be emphasized that "[ADM systems that draw on some automated learning] can generally vary in the complexity of how inputs determine outputs—including simple threshold rules for single input variables—, as well as in the extent to which humans are involved in the final decisions" [Ge22]. The more complex ADM systems are, the more difficult it can be to understand their decision logic as well as results – especially for the case of ML-based systems. Against the background of decreasing resources in the public sector, experts agree that algorithmic systems are increasingly used [Bu20; Hi18]. However, the use of these tools is also accompanied by risks of discrimination, which we discuss in the next section.

## 2.2 Algorithmic gender discrimination in the public sector

In general, discrimination can be understood as the unequal treatment of persons on the basis of protected (personal) characteristic such as their gender [Or20; KW20]. From a social science perspective, discrimination is "the categorization of different social groups with different social positions" (own translation) [KO20] that can be used as a justification for disadvantageous treatment. "Discrimination is consequently a social practice that restricts access to certain material as well as immaterial goods on the basis of (perceived) group membership" (own translation) [KO20]. A distinction is made between *direct discrimination*, i.e., the treatment of a person depends directly on protected characteristics and indirect discrimination, i.e., the treatment of an individual is not directly related to protected characteristics, but is linked with correlation to it [KO20; Or20]. *Indirect discrimination* can also be referred to as *systematic* or *unintended discrimination* [Or20]: Because algorithmic discrimination can run unintentionally via correlations to protected characteristics the detection of discrimination can be extremely difficult [Or20].

The use of algorithmic systems in decision-making processes can reinforce (indirect) discrimination [Ma21]. One of the possible reasons for this is that the ADM systems might operate on the basis of biased data. Lopez (2021), Allhutter and colleagues (2020) and Gerdon and colleagues (2022) provide insights into different types of data biases in software systems: Selective participation and representation of social groups as well as other technical or conceptual mismeasurements in the data generation and collection process can lead to a misinterpretation of reality (*technical bias*) or a deviation of the data basis from the phenomenon that is to be represented (*sociotechnical bias*). In contrast to that, *social biases* represent given values embedded in society (also referred to as *historical discrimination*). However, biases in ADM systems not only stem from the data generation and collection process. Gerdon and colleagues (2022) and Smith/Rustagi (2021) emphasize that biases can also be produced during data preparation and analysis due to wrong decisions regarding the labelling and processing of data (modeling). If a machine learning model is deployed, *indirect discrimination* can be further reinforced by applying whole statistical models instead of substitute variables, which include a multitude of variables and their weighted relations [Or20].

# 3 Gender-relevant use cases from the public sector

## 3.1 Austria: Discrimination by whom?

In Austria, the PES[5] introduced an ADM system that calculated the probability of unemployed persons to get a job within a certain time period *(labor market integration chance)* [HW18; Or20]. The chance was calculated for the registered unemployed on the basis of variables such as age, education and previous occupation. In the Austrian model, a chance above 66% means that a person has a high chance getting a job (group A[6]). If the chance is between 25 and 66% a person is considered to have medium chances (group B), and a chance lower than 25% indicated low chances in the long-term perspective (group C[7]). As job seekers within group A were believed to find employment without further assistance, and job seekers within group C would need different support, a majority of the resources for employment-supporting measures was allocated to group B. As the variables ‘gender’ and ‘care responsibilities’ had a negative impact on the score, women were overrepresented in group B, where they received further assistance. But women in comparison to men with similar employment histories were also twice as likely to be assigned to group C, thus with the lowest calculated chances of labor market reintegration [Al20]. According to the PES, this reflected the reality of the labor market [Al20; Al21; HW18]. The head of the PES Austria emphasized that due to the disproportionate representation of women in group B, women would also receive disproportionately more benefits and this could compensate for the disadvantaged position of women in the labor market [FD19].

The key question here is whether the unequal classification of men and women by the ADM system is already discriminatory [KO20]. Direct discrimination on the basis of gender does not apply here. Nevertheless, it results in *indirect discrimination* as the algorithm assigns a certain (negative) weighting and thus a lower score based on protected characteristics such as gender [KO20]. The same applies to any ‘care responsibilities’, a variable that only has a negative impact on the score of women. The developers argue that care responsibilities statistically do not have a negative impact on men's labor market opportunities [FD19]. Allhutter and colleagues (2019), on the other hand, emphasize that the ADM system is not a true representation of the labor market,

[5] Equivalent to the ‘Bundesagentur für Arbeit’ in Germany.
[6] Persons in group A have a 66% probability of being employed for at least 90 days in the next seven months.
[7] For persons in this group the probability of being employed for at least 180 days in the next 24 months is less than 25%.

but is distorted in many ways. The deployment of a lower weighting for women would lead to statistical discrimination [FD19]. This would also have an amplifying effect if several characteristics apply to job seekers, such as being a woman and having a migration background. Members of this group are disproportionately found in the group with the lowest labor market opportunities, in which Allhutter and colleagues (2019) see the danger of self-reinforcing processes because access to cost-intensive resources of the middle segment is systematically impeded for this group (*cumulative disadvantage*) [Al20; Or20]. Differentiation by gender may entrench existing problems [Or20]. This is especially crucial in decisions about elementary basic needs such as work and social welfare, where the risk of discrimination must be considered during the first stages of data processing [Or20]. Even the statistical mapping of discriminatory structures can become a self-fulfilling prophecy and reinforce discrimination. Therefore, Fröhlich and colleagues (2018) argue that it is not the algorithm that discriminates, but the institution that uses it. In 2020, the Data Protection Authority in Austria prohibited the use of the ADM system, which was overturned by the Federal Administrative Court. The case is currently pending before the Administrative Court.

### 3.2 Poland: Same same but different?

Similar to Austria, the PES in Poland used an ADM system to calculate the labor market chances of job seekers, dividing them into three categories that are each supported through different labor market programs [Or20]. The system's calculation is based not only on the socio-demographic data of the job seekers but also on information provided by these persons in interviews. Group 1 included those unemployed persons who, with less than 22 minus points, were considered to have a high probability of quickly finding a job again. Group 2 included unemployed persons who had between 23 and 59 minus points. The 'third profiled' group included all those who had more than 59 minus points and were thus considered to have a low chance of finding a job [Ni19]. In most employment offices in Poland, approximately equal numbers of women and men were assigned to each group. In some employment agencies, however, the data showed that women were predominantly (60-70%) represented in group C, thus with the assumed lowest chances of labor market reintegration [Al20; Al21]. Because the gender of women influenced their calculated labor market chance negatively, there is a risk of *direct discrimination* [Or20]. *Indirect discrimination* is assumed to exist in this example due to the application of the characteristics ‘times for childcare and care’, which statistically affect women more often than men [Or20; Ni15]. Niklas and colleagues (2015) explain that the mere fact that these variables influence the labor market profile of job seekers can lead to unequal treatment of women. Accusations of discrimination came not only from women, but also from other groups such as single mothers [Ni19]. Due to concerns about discrimination and the ineffectiveness of the system, Poland's Constitutional Court ruled that the ADM system violated the constitution. Today, the system is no longer in use [Ni19].

# 4 Three lessons from the private sector

The use cases from Austria and Poland outlined in chapter three demonstrate how the deployment of ADM in the public sector can affect gender equality. Meanwhile, ADM deployment in the private sector already provides a plethora of gender-relevant use cases and examples, ranging from systematically discriminating hiring algorithms [Da18] to racist and sexist chatbots [TC16], to unfairly targeting job ads [Ho21]. As the number of cases rises, so do the critical voices that point out risks and shortcomings of the systems. Scholars of various disciplines such as ethics, social sciences, law or computer sciences provide insights on the risks of algorithmic bias [e.g., Ma21; WH22; Ad22]. While the issue of algorithmic gender bias is by no means solved in the private sector, there are valuable lessons to be learned on how the public sector might mitigate those emerging risks. In the following, we identify three key lessons based on recent developments. They do, however, not present an exhaustive list.

## 4.1 Framing ADM as fallible and empowering civil servants to challenge it

Both unconscious and conscious gender biases are omnipresent in society. As emphasized above, those gender biases are inherently incorporated in every recommendation or decision made by ADM systems [Ke20]. Popular examples of gender discrimination are no longer exclusively discussed in academic circles, but obtain high media coverage and are part of the public discourse [e.g., Da18; TC16; Ho21]. However, even as society is becoming more aware of the risks, the available education on algorithmic gender bias is currently not sufficient. ADM is increasingly incorporated in human resource management, e.g., to provide recommendations on which candidates to invite for an interview, which employees to develop, and whose productivity to closely monitor [Tu18]. These systems, often summarized as *people analytics*, continuously analyze behavioral data of employees and promise tempting benefits, also to employees themselves. E.g., they can improve self-organization by offering personalized feedback on individual work patterns or predict burnout risks [EQ18]. Presented with the benefits of ADM, employees might not be aware that those come at a price. The constant data collection of the systems can, e.g., invade the (data) privacy of employees [Gi21]. In the private sector, works councils and unions have proven as powerful means to provide education on these risks and mitigate them on behalf of employees [Ml21]. Low-threshold education is vital so that employees can protect themselves from potential harm. However, with increasing technical complexity, e.g. due to elements of machine learning, it also becomes increasingly challenging to trace and understand how recommendations are derived. In the Polish PES, civil servants working with the ADM system reported fear of questioning or resisting the automated decision recommendation [Ni15]. This can be attributed to automation bias, i.e. humans favoring automated decisions over those made by humans as the technology is thought to be less likely to fail [PM10]. To prevent this, algorithms should be framed as fallible, imperfect companions in organizations according to a recent study on *people analytics*

[Ga20]. One should focus on the actual benefits of algorithms to efficiently process large amounts of data rather than initiating ever new discussions about algorithms as perfect and rational decision-makers, or even human-like sentient beings [Jo22]. High but oftentimes unattainable expectations of ADM prevent users from grasping its very real risks. Daring to challenge and question automated decisions requires psychological safety and an absence of fear in public administration. The opacity of an ADM system which was used by the PES in Portugal led to mistrust among civil servants leading them to refocus on their own decision-making competencies [Ze20]. To empower and encourage civil servants to critically evaluate and challenge decisions made by ADM systems *before* its implementation is as important as the education about potential risks of algorithmic gender bias.

### 4.2 Addressing power imbalances as root causes of algorithmic gender discrimination

Gender-based discrimination does not only occur in the workplace, but now permeates many essential areas of life. In finance and insurance, assigning scores and measuring creditworthiness based on certain metrics such as employment history has long been a reality. This assessment is increasingly automated using ADM systems. When Apple issued its own credit card in 2019, numerous women complained about the credit limit assigned to them, if they were granted credit at all [Mo20]. It was particularly noticeable how credit lines were given to people in marriages. Spouses who had joint accounts and thus the same financial resources received significantly different scores. While men received scores matching their actual resources, women were often rejected or received lower amounts [Vi19]. This is because of the deployment of historical data: In the past, women were usually not able or allowed to work full time and therefore did not provide significant income for their families. The historical data sets used by credit card companies such as Apple to calculate these scores still contain traditional gender biases, and thus still suggest that women are not capable of paying back a loan due to their assigned role in traditional family-settings [Ca21]. When training the models, programmers should have realized that they are handling highly biased data. In this case, and in many others, they did not recognise that the data was in favor of men. One cause for this is seen in the lack of diversity amongst the teams who work on data models. As they mostly consist of white young men, they might simply not recognize when, for example, data input for women with their own stable income, is missing. It is vital that women and other marginalised or vulnerable groups are included in the developing processes of ADM, as the results of the decision making concern the whole society. A key issue that perpetuates human bias in algorithms is the distribution of power and hierarchies in the workplace. A recent study by Miceli and colleagues (2021) examines how exactly biases are reenacted by algorithms. Acknowledging that historically evolved training datasets almost always contain human biases, the study examines how potentially biased labels and datasets are dealt with in the programming process. One of their elementary findings is that power in the form of seniority and financial resources plays a crucial role. Customers who purchase an ADM system and supervisors of teams that develop them often hold the power to decide which labels or which data will be used in the final version of the system. Usually neither of those two groups is likely to be skilled in machine learning, advanced statistics or data science and their decisions and preferences might lead to inaccurate and biased systems. Public servants often have a similar skills gap. When ADM is implemented, they not only have to understand how the systems are functioning, but they have to be aware of unconscious biases. Thus, decision-makers in public sector agencies could adjust the allocation of training resources for their employees accordingly. Mandatory unconscious bias trainings could potentially raise this awareness and prevent wrongful decisions by public administration. Regulation is another important lever to counter power imbalances. Krafft and colleagues (2022), for example, propose a risk-based approach similar to that of the proposed AI Act of the European Union.

### 4.3 Embracing criticism and responding to whistleblowing – lessons on accountability

Money and power play a role not only in the development of ADM systems, but also in the way its results are discussed, disseminated, and explained. Google is a pioneer in the technology sector, and approaches to data-based employee management in particular have been developed there. At the same time, Google has also made negative headlines, for example for unresolved issues of perpetuating harmful gender stereotypes in its image search and text-to-image diffusion model *Imagen* [Di15; Me22]. When Timnit Gebru, a member of Google's in-house ethics team, collaborated with other female researchers for a paper on the potential risks of racial discrimination in language models, she seemingly crossed an invisible line [Be21]. Google demanded that she either withdraw the paper or remove her name from it. In subsequent discussions, Gebru was dismissed by Google. She is considered a whistleblower by many scientists. The paper in question was about uncovering risks so that they can be mitigated and become part of scientific discourse. It was the harsh treatment of her that caused many headlines and brought the case into the public discourse. For many, it left the impression that Google feared the results of Gebru's study. In the workplace – and in academia in particular – employees are highly dependent on those in power at their organization. A dismissal by a company like Google could be an existential threat to many scientists. If they refrain from conducting meaningful research on ADM out of fear for personal consequences, this will have far-reaching consequences. The Polish example, where civil servants working with the ADM system reported fear of questioning the automated decisions [Ni15], underlines the importance of public administration to embrace and welcome criticism to continuously improve its ADM systems. Studies have found that algorithmic discrimination causes less moral outrage than human discrimination [Bi22], thus it is ever more crucial to listen to those that speak up.

### 4.4 Mitigating gender discrimination as a shared task

Who is responsible for driving gender-sensitive automation in the public sector? The lessons from the private sector are highly intertwined. They are not aimed solely at policymakers and heads of public agencies, but are intended to appeal to everyone involved with ADM systems in the public sector. In the following, we outline selected actions on how key public sector actors can contribute to gender-sensitive algorithmic automation:

1. *Civil servants:* Civil servants need to be empowered to critically evaluate the decision-making proposals of ADM systems – especially with regard to discrimination risks. Therefore, technical training should be integrated into the career training and degree programs of civil servants.
2. *Process designers:* Whether technology will have its intended effect depends on how it is socially embedded in the specific work context. In addition to technical expertise, case manager need to be granted sufficient time to review automated decision proposals (for non-intended, discriminatory effects).
3. *Management:* If civil servants notice that the use of ADM systematically discriminates against groups of people, such as women, it should be possible to establish transparency quickly. If ADM systems are not modified or discontinued as a result, whistleblowers should be effectively protected from personal consequences.
4. *Policymakers:* The use of ADM systems is often mentioned politically in the same breath as efficiency (such as shorter waiting times for citizens and cost savings). However, the goals of technology use, especially in decision-making, can also be qualitative, such as fundamentally questioning (and modernizing) decision-making bases and processes. This can be politically demanded and promoted.

Since the public sector should be a role model when it comes to adopting technology responsibly, we hope that these lessons provide valuable insights on how to develop more gender-responsive ADM systems for the public sector. Despite all efforts to learn from past experience, universal guidelines or standardized technical procedures to prevent discrimination do not exist. Instead, they must be based on the specific application. After all, the concept of fairness is not globally agreed on either [KO20].

## 5 Conclusion

In this paper, we evaluated ADM and gender discrimination in the context of the public sector. Using the examples of ADM systems in the Public Employment Services in Austria and Poland, we demonstrated potential risks of ADM systems for gender equality. The usage of historical data oftentimes provides no accurate measurement for the reality of women today (the Corona pandemic, at the latest, made it clear once again how distorted and unrepresentative historical data can be). In fact, it was not until 1958 that women in Germany were legally allowed to hold their own bank accounts without their husband's consent [Bu18]. Financial matters of women were often controlled by the men in their lives, such as fathers and husbands. This historic discrimination is still reflected in the many (unconscious) biases women encounter in their lives today. The public sector is legally required to offer discrimination-free service to all civilians. If ADM systems are used to distribute services to civilians, it has to be ensured that they entail no risks of discrimination. However, mitigating these risks is no easy task. By turning to the public sector, we identified three core lessons to keep in mind when developing such technology for the public sector: framing ADM as fallible, addressing power imbalances and protecting whistleblowers. Above all, the challenges and consequences of ADM as described in our paper show that ADM is always political. Just as algorithms inherit biases against women, they have a strong potential to perpetuate discrimination based on religion, disabilities, social background, origin and many other biases found in western societies. The stakes are high: Citizens rely on fair access to and distribution of social welfare and support. Therefore, more research is needed on the drastic real-world consequences that might entail the deployment of ADM systems without mitigating the risks. The issue of gender discrimination is fundamentally inherent in algorithmic systems [La16], thus the conversations of mitigating it must be continuously ongoing.

## Bibliography


[Ad22] Adams-Prassl, J.: Regulating Algorithms at Work: Lessons for a 'European 'European Approach to Artificial Intelligence'. In: European Labour Law Journal (13:1), London: SAGE Publications, 2022.

[AW19] AlgorithmWatch, AW AlgorithmWatch gGmbH: Atlas of Automation – Automated decision-making and participation in Germany. 2019. https://atlas.algorithmwatch.org/wp-content/uploads/2019/04/Atlas_of_Automation_by_AlgorithmWatch.pdf, accessed: 9/7/2022, 2019.

[AW20] AlgorithmWatch, AW AlgorithmWatch gGmbH: Automating Society Report 2020. 2020. https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/12/Automating-Society-Report-2020.pdf, accessed: 9/7/2022.

[Al20] Allhutter, D.; Mager, A.; Cech, F.; Fischer, F.; Grill, G.: Der AMS-Algorithmus. Eine Soziotechnische Analyse des Arbeitsmarktchancen-Assistenz-Systems (AMAS). Endbericht. 2020. In (Austrian Academy of Sciences, Institute of Technology Assessment): ITA-Projektbericht, nb. 2020-02, 2022.

[Al21] Allhutter, D.: Ein Algorithmus zur effizienten Förderung der Chancen auf dem Arbeitsmarkt? In (Institute for Social and Economic Sciences): Wirtschafts- und Sozialpolitische Zeitschrift (WISO), Nr. 1, 2021.

[Be21] Bender, E.M.; Gebru, T.; McMillan-Major, A.; Shmitchell, S.: On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 610-623, 2021.

[Bi22] Bigman, Y.; Gray, K.; Waytz, A.; Arnestad, M.; Wilson, D. (2022). Algorithmic Discrimination Causes Less Moral Outrage than Human Discrimination. In: Journal of Experimental Psychology Gen., 2022.

[Bu18] Gleichberechtigung wird Gesetz. In: Federal Agency for Civic Education (Online), 27 June 2018, https://www.bpb.de/kurz-knapp/hintergrund-aktuell/271712/gleichberechtigung-wird-gesetz/, accessed: 10/7/2022.

[Bu20] Busuioc, M.: Accountable Artificial Intelligence: Holding Algorithms to Account. Public Administration Review, Vol. 81(5), pp. 825–836, 2020.

[Ca21] Campbell, I.C.; The Apple Card doesn't actually discriminate against women, investigators say. In: The Verge (Online), 23 March 2021, https://www.theverge.com/2021/3/23/22347127/goldman-sachs-apple-card-no-gender-discrimination, accessed: 07/7/2022.

[Da18] Dastin, J.: Amazon scraps secret AI recruiting tool that showed bias against women. In: Reuters (Online), 11 October 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G, accessed: 9/7/2022.

[Di15] Dishman, L.: The Hidden Gender Bias In Google Image Search. Fast Company (Online), 22 April 2015, https://www.fastcompany.com/3045295/the-hidden-gender-bias-in-google-image-search, accessed: 9/7/2022.

[El21] Elyounes, D.A.: 'Computer Says No!': The Impact of Automation on the Discretionary Power of Public Officers. In: Vanderbilt Journal of Entertainment and Technology Law, Vol. 23 (3), pp.450-515, 2021.

[EQ18] Estévez-Mujica, C. P.; Quintane, E.: E-mail communication patterns and job burnout. In: PLOS ONE, Vol. 13(3), 2018.

[FD19] Fröhlich, W.; Spiecker named Döhmann, I.: Können Algorithmen diskriminieren? djbZ, Zeitschrift des Deutschen Juristinnenbundes Vol. 22 (2), pp. 91-93, Baden-Baden: Nomos, 2019.

[Ga20] Gal, U.; Jensen, T. B.; Stein, M.-K.: Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. In: Information and Organization, Vol. 30(2), 2020.

[Ga22] Gardner, A.: Responsibility, Recourse, and Redress: A Focus on the Three R's of AI Ethics. Commentary. In: IEEE Technology and Society Magazine Vol. 41 (2), 2022.

[Ge22] Gerdon, F.; Bach, R.L.; Kern, C.; Kreuter, F.: Social impacts of algorithmic decision- making: A research agenda for the social sciences. In: Big Data & Society, Vol. 9(1), pp. 1-13. London: SAGE Publications, 2022.

[Gi21] Giermindl, L. M.; Strich, F.; Christ, O.; Leicht-Deobald, U.; Redzepi, A.: The dark sides of people analytics: Reviewing the perils for organisations and employees. In: European Journal of Information Systems, Vol. 31(3), pp. 410-435, 2021.

[Hi18] Hill, H.: Was bedeutet Künstliche Intelligenz (KI) für die Öffentliche Verwaltung? Verwaltung und Management, Vol. 24(6), pp. 287-294, 2018.

[HW18] Holl J.; Kernbeiß, G.; Wagner-Pinter, M.: Das AMS-Arbeitsmarktchancen-Modell. Dokumention zur Methode. Konzeptunterlage. In (Synthesisforschung, ed), https://www.ams-forschungsnetzwerk.at/downloadpub/arbeitsmarktchancen_methode_%20dokumentation.pdf, accessed: 9/7/2022. Wien, 2018.

[Ho21] Horwitz, J.: Facebook Algorithm Shows Gender Bias in Job Ads, Study Finds. In: The Wall Street Journal (Online), 9 April 2021.

[HT20] Hütt, M.T.; Schubert, C.: Fairness von KI-Algorithmen. In (Mainzer, K., ed.): Philosophisches Handbuch Künstliche Intelligenz. Wiesbaden: Springer, 2020.

[Jo22] Johnson, K.: LaMDA and the Sentient AI Trap. In: Wired (Online), 14 June 2022, https://www.wired.com/story/lamda-sentient-ai-bias-google-blake-lemoine/, accessed: 9/7/2022.

[Ke20] Kellogg, K. C.; Valentine, M. A.; Christin, A.: Algorithms at Work: The New Contested Terrain of Control. In: Academy of Management Annals, Vol. 14(1), pp. 366–410, 2020.

[KO20] Kolleck, A.; Orwat, C.: Mögliche Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick. In (Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, ed.): Hintergrundpapier Nr. 24, 2020.

[KW20] Köchling, A.; Wehner, M.C.: Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making the context of HR recruitment and HR development. In: Business Research, Vol. 13, pp. 795–848, 2020.

[Ke20] Krafft, T.; Zweig, K.A.; König, P.: How to regulate algorithmic decision-making: A framework of regulatory requirements for different applications. In: Regulation & Governance, Vol. 16 (1), Melbourne: John Wiley & Sons Australia, 2020.

[La16] Laaff, M.: Kate Crawford über KI: ‚Diese Systeme sind fürs Diskriminieren designt worden', 2022. In: ZEIT (Online), 8 June 2022, https://www.zeit.de/digital/internet/2022-06/kate-crawford-ki-machine-learning-forschung, accessed: 9/7/2022.

[Le21] Levi, K.; Chasalow, K.E.; Riley, S.: Algorithms and Decision-Making in the Public Sector. In: Annual Review of Law and Social Science, Vol. 17; pp. 309–334, 2021.

[Lo21] Lopez, P.: Diskriminierung durch Data Bias Künstliche Intelligenz kann soziale Ungleichheiten verstärken. In (WZB, ed.): WZB Mitteilungen, Vol. 171, pp. 26-28, 2021.

[Ma21] Marjanovic, O.; Cecez-Kecmanovic, D.; Vidgen, R.: Algorithmic Pollution: Making the Invisible Visible. In: Journal of Information Technology. London: SAGE Publications, 2021.

[Me22] Metz, R.: AI made these stunning images. Here's why experts are worried. CNN Business (Online), 1 July 2022, https://edition.cnn.com/2022/06/30/tech/openai-google-realistic-images-bias/index.html, accessed: 9/7/2022.

[Mi22] Miceli, M.; Posada, J.; Yang, T.: Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power? In (Association for Computing Machinery, ed.): Proceedings of the ACM on Human-Computer Interaction (forthcoming), Vol. 6, New York, pp. 1–14, 2022.

[Ml21] Loi, M.: People Analytics muss den Menschen zugutekommen: Eine ethische Analyse datengesteuerter algorithmischer Systeme im Personalmanagement (No. 450). Study der Hans-Böckler-Stiftung, 2021.

[Mo20] Mozannar, H.; Ohannessian, M.; Srebro, N.: Fair learning with private demographic data. In: Proceedings of the 37th International Conference on Machine Learning, 119, pp. 7066-7075, 2020.

[Ni15] Niklas, J.; Sztandar-Sztanderska, K.; Szymielewicz, K.: Profiling the Unemployed in Poland: Social and Political Implications of Algorithmic Decision Making. Fundacja Panoptykon, ed., Warsaw, 2015.

[Ni19] Niklas, J.: Poland. Algorithms Deciding the Future of Social Rights: Some Early Lessons from Poland. In (Global Information Society Watch, ed.): Artificial Intelligence. Human rights, social justice and development, pp. 195-200, 2019.

[Or20] Orwat, C.: Risk of Discrimination through the Use of Algorithms. Federal Anti-Discrimination Agency, ed., first edition (German) 2019, 2020.

[PM10] Parasuraman, R.; Manzey, D.H.: Complacency and Bias in Human Use of Automation: An Attentional Integration. In: The Journal of the Human Factors and Ergonomics Society, Vol. 52(3), pp. 381-410, Los Angeles: SAGE, 2010.

[SU22] Schwarting, R.; Ulbricht, L.: Why Organization Matters in ‘Algorithmic Discrimination’. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie (KZfSS), Vol. 75, pp. 307–330, 2022.

[SR20] Smith, G.; Rustagi, I.: Mitigating Bias in Artificial Intelligence. An Equity Fluent Leadership Playbook. University of California (Berkeley), Center for Equity, Gender & Leadership, 2020.

[Sp20] Spiekermann, S.: Zum Unterschied zwischen künstlicher und menschlicher Intelligenz und den ethischen Implikationen der Verwechselung. In (Mainzer, K., ed.): Philosophisches Handbuch Künstliche Intelligenz. Wiesbaden: Springer, 2020.

[TC16] Tennery, A.; Cherelus, G.: Microsoft's AI Twitter bot goes dark after racist, sexist tweets. In: Reuters (Online), 24 March 2016, https://www.reuters.com/article/us-microsoft-twitter-bot-idUSKCN0WQ2LA, accessed: 9/7/2022.

[Tu18] Tursunbayeva, A.; Di Lauro, S.; Pagliari, C.: People analytics - A scoping review of conceptual boundaries and value propositions. In: International Journal of Information Management, Vol. 43, pp. 224–247, 2018.

[Vi19] Vigdor, N.: Apple Card Investiagted After Gender Discrimination Complaints. In: New York Times (Online), 10 November 2019, https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html, accessed: 07/7/2022

[WH22] Weiskopf, R.; Hansen, H. K.: Algorithmic Governmentality and the Space of Ethics: Examples from ‘People Analytics’. In: Human Relations, pp. 1–24. London: SAGE Publications, 2022.

[Ze20] Zejnilović, L.; Lavado, S.; Martínez de Rituerto de Troya, I.; Sim, S.; Bell, A.: Algorithmic Long-Term Unemployment Risk Assessment in Use: Counselors’ Perceptions and Use Practices. In: Global Perspectives, Communication and Media, 2020.