# Reinforcing Intersectional Inequality via the AMS Algorithm in Austria

Paola Lopez[1]

[1]Faculty of Mathematics, University of Vienna

**Abstract.** This paper examines the so-called *AMS Algorithm* from a mathematical perspective: this algorithmic system constitutes a predictive model that will be used by the Public Employment Service Austria (AMS) starting in 2020 to algorithmically classify job-seekers into three groups, each with different access to AMS support resources, according to their predicted *chances* on the labour market. Since the features gender, age, childcare responsibilities, disability and citizenship are *explicitly implemented* in the model and are thus linked to the availability of resources, this algorithmic system is to be considered very problematic. This paper is part of an ongoing research project, and it identifies three conceptual building blocks of the AMS Algorithm that are all based on human decisions and in which obvious societal bias can be located. Furthermore, this model is used as an illustrative example to address the larger question of what can be expected when predictions are made that are based solely on data that describes the past: If the predictions by these models result in unquestioned and confirmatory measures such as the redistribution of resources, a reproduction and reinforcement of inequality is possible. If these measures are now applied to vulnerable and highly dependent target groups, such as job-seekers, it will be more drastic: In a first step, these predictive models depict the reality of discrimination, then, in a second step, normatively reinforce it as a supposedly objective fact and finally, in a third step, return it to the social sphere by means of the resulting measures.

## 1 Introduction

Starting in 2020, the Public Employment Service Austria (Arbeitsmarktservice Österreich, in short AMS) will use a predictive model (*Arbeitsmarkt-Chancen-Modell*) to segregate job-seekers into groups with different access to AMS support resources according to their predicted *chances* on the job market. It became known in the media by the name *AMS Algorithm* primarily through the publication of its accompanying method paper. It can be inferred from the paper that the personal data entry *Gender: Female* results in an automatic deduction of points, which means that a woman can be assigned to a group with less access to AMS resources solely on the basis of her gender. Further potential point deductions according to personal data, such as age or nationality, can lead to an intersectionally compounded disadvantage: Figure 1 below shows a screenshot taken from the method paper that was discussed widely in the media.

```
BE_INT
 = f ( 0,10
      – 0,14 x GESCHLECHT_WEIBLICH
      – 0,13 x ALTERSGRUPPE_30_49
      – 0,70 x ALTERSGRUPPE_50_PLUS
      + 0,16 x STAATENGRUPPE_EU
      – 0,05 x STAATENGRUPPE_DRITT
      + 0,28 x AUSBILDUNG_LEHRE
      + 0,01 x AUSBILDUNG_MATURA_PLUS
      – 0,15 x BETREUUNGSPFLICHTIG
      – 0,34 x RGS_TYP_2
      – 0,18 x RGS_TYP_3
      – 0,83 x RGS_TYP_4
      – 0,82 x RGS_TYP_5
      – 0,67 x BEEINTRÄCHTIGT
      + 0,17 x BERUFSGRUPPE_PRODUKTION
      – 0,74 x BESCHÄFTIGUNGSTAGE_WENIG
      + 0,65 x FREQUENZ_GESCHÄFTSFALL_1
      + 1,19 x FREQUENZ_GESCHÄFTSFALL_2
      + 1,98 x FREQUENZ_GESCHÄFTSFALL_3_PLUS
      – 0,80 x GESCHÄFTSFALL_LANG
      – 0,57 x MN_TEILNAHME_1
      – 0,21 x MN_TEILNAHME_2
      – 0,43 x MN_TEILNAHME_3)
```

**Fig. 1. Published $f_1$-coefficients for the base population, screenshot from the method paper (Holl, et al., 2018, p. 11)**

## 2 Classification

The *AMS Algorithm* uses different types of data (see below) to model the probabilities of the job-seekers to achieve two *goals*, namely the *short-term goal* $f_1$ and the *long-term goal* $f_3$, see below. Using these probabilities, three groups of job-seekers are formed: (Holl, et al., 2018)

- **Group A**: Job-seekers who are predicted to achieve the short-term goal with a probability $f_1$ of at least 66% are said to have *high chances* on the job market, according to the model. Therefore, they are less eligible for support from the

AMS resources since they are not considered to need much support based on their predicted already high chances. (Kopf, 2018a)

- **Group C**: Job-seekers who will achieve the long-term goal with a predicted probability $f_3$ of less than 25% are classified as having *low chances* according to the model. They should get access to *different* resources in order to prevent *expensive resources* to be used on people with *little output* (Kopf, 2018a): external support formats have been tested in a pilot project in 2018, see below.
- **Group B**: Those job-seekers who fall neither into Group A, nor into Group C, are said to have *medium chances* on the labour market. The AMS plans to focus on this group of job-seekers, they should get full access to the AMS resources. (Kopf, 2018a)

## 3 The Model

It is therefore essential for job-seekers which of the three groups they are assigned to, hence the prediction model itself will be examined next in order to assess which factors have how much influence on the resources available to jobseekers according to this triage classification.

### 3.1 Base Population

According to the published method paper, three types of data are relevant for the calculated probabilities $f_1$ and $f_3$, firstly so-called *personal features*, secondly the previous individual employment history, and thirdly the current AMS case. (Holl, et al., 2018, p. 3) Furthermore, a differentiation must be made as to which *model variant* is used in each case - job seekers are first divided into different statistical *populations* (i.e. subgroups of job seekers) with regard to the *quality of information* (i.e. data) available in the respective case, so that a different *model variant* is realized for each *population*. In this context, good *quality of information*, which defines the so-called *base population* of job-seekers, refers to the availability of *continuous* data and employment history with social security status in the previous 48 months. According to the method paper, the calculation of the probabilities $f_1$ and $f_3$ is most possible for the base population. (Holl, et al., 2018, p. 4) This corresponds to the statistical principle that predictions are possible with greater accuracy, the more *relevant* information is available and can thus be incorporated into the predictive model. (Hastie, et al., 2008) If less data is available from the past, the method paper speaks of subpopulations which, due to the lack of important data, cannot be estimated as well as the base population. (Harrell, 2015)

### 3.2 The Coefficients

The probabilities for the *short-term goal* $f_1$ and the *long-term goal* $f_3$ are calculated using logistic regression. (Holl, et al., 2018, p. 7) This means that the model, or more precisely, each *model variant*, is determined by a list of weights of features, i.e. positive or negative coefficients (numbers) that describe the positive or negative influence of different features on the probabilities. The key to understanding the classification of job-seekers into groups is therefore, on the one hand, the features that are included and, on the other hand, the respective weights (coefficients) of these features. As mentioned above, a type of features is that of the so-called *personal features*. These include age, gender, nationality, education, childcare responsibilities and disability. (Holl, et al., 2018) Due to the limited scope, this paper focuses on the personal features, as these are among the explicitly protected features by legal anti-discrimination regulations. (Holzleithner, 2016)

The coefficients for the *short-term goal* $f_1$ of the model variant for the base population were published as example, see Fig. 1. The negativity/positivity of the coefficients for the different features appear here as an (undoubtedly unplanned) intersectional decoding of social inequalities: The coefficients for the data entry *Gender: Female*, for being of an age above 30, for non-EU citizenship and for disability are negative meaning that these features negatively influence the probability of reaching the *short-term goal.* Childcare responsibilities also have a negative coefficient, which, however, is only taken into account if the individual is female, reflecting the statistical finding that having childcare responsibilities does affect women's probability $f_1$ of job placement, but not men's. (Kopf, 2018b)

The coefficients were determined by analysing the available data from the past on the basis of the two *goals.* Since it is known in retrospect exactly which persons achieved which goals, the coefficients that encode the impact of each feature on job placement in the past could be estimated. These are used to make predictions about the future.

## 4 Three Building Blocks

From a mathematical perspective, the AMS Algorithm, as well as any other such classification system using logistic regression in an equivalent manner, consists of three basic building blocks: the data, the target variables, and the thresholds. Hence, the coefficients as well as the distribution of all job-seekers to the Groups A, B and C depends on these three components. In the following chapter it will be discussed how any change in one of these building blocks would lead to a different decomposition of the job-seekers, so that the categorization as it is cannot be regarded as a given and neutral one, but in this sense has a certain degree of fragility to it: After all, crucial parts of the building blocks are based on human decisions, which, as will be detailed below, could also have turned out differently.

Each of the three building blocks has a conceptual dimension, as well as a concretely implemented dimension: The (specific view on the) past as a conceptual dimension finds its specific realization in the data; the (particular outlook on) the future is implemented as the target variables, or *the goals*; and the thresholds are the numerical cut-off points that describe the valuations that stand behind the decomposition of the job-seekers.

### 4.1 The (Specific View on the) Past = The Data

The available data determines both the categories of features that can be statistically (and thus algorithmically) analysed, (Hong, 2016) and the resulting coefficients of the model. (Zheng & Casari, 2018) The data that was used to build the model is data that the AMS has been collecting and evaluating for a long time. The types of data that the AMS is legally allowed to collect are stated in § 25 AMSG, and include personal data, as well as employment history data. The data used was therefore not collected for the purpose of developing the model: Existing data was used to find statistically significant correlations between available features and job placement rates.

The data (in Machine Learning one speaks of *training data* that is used to *train* a model to make correct predictions (Goodfellow, et al., 2016)) is always essential for the model that is to be developed. The probabilities from which the coefficients that form the core of the model are determined on the basis of the data. After assuming a logistic regression-approach and by fitting the model using the maximum likelihood method (which basically estimates coefficients by maximizing plausibility (Hastie, et al., 2008)), the respective impact of the different features is estimated, leading to this degree of impact being reflected in the coefficients. *The available data therefore constitutes the past on the basis of which the future is to be predicted via the model.*

The question that has been investigated is: *Which categories* of people have successfully achieved job placement, *when* and for *how long*? The solution to this question is available within the AMS-internal data concerning past cases, and so the probability of achieving the *short-term goal* and the *long-term goal* was retrospectively assessed. This data – personal data, data on the previous employment history, and AMS-internal case data – thus reflects (to the extent of the model assumptions) how the labour market has reacted to job-seekers that are recorded within the AMS data in the past.

If a person is not sufficiently datafied, for example if the person's employment history is *fragmented* (see below), then there are gaps in the data. Missing data is a common problem in Machine Learning (Harrell, 2015), and here it was dealt with by developing other model variants for these people (see above). The way in which this was done, or more detailed information on the strategies for dealing with data gaps, cannot be found in the method paper.

### 4.2 The (Particular Outlook on the) Future = The Two Goals

The particular outlook on the future refers to the events for which the probability is calculated by the predictive model. In general, a probability can only be estimated with respect to a specific and very concretely defined event that is quantifiable and included in the training data. It is thus essential to examine the concrete definition of the *short-term goal* and the *long-term goal*.

- The *short-term goal*, which is relevant for $f_1$, is defined as *successful* if the job-seeker in question achieves job placement for at least three months (90 days) within the next seven months. (Holl, et al., 2018, p. 7)
- The *long-term goal* for $f_3$ is defined as successful if job placement for at least six months (180 days) is achieved within the next 24 months. (Holl, et al., 2018, p. 7)

The probabilities are computed accordingly. For example, an output of $f_1$ = 0,59 means that, according to the model, the person is predicted to have a 59% chance of achieving job placement for three months or more within the next seven months, since persons with the same data entries achieved this same goal with a probability of 59% in the observed past data.

The method paper did not provide explanatory reasons for these concrete definitions, nor is it explained whether these timeframes and job placement goals are AMS-internal objectives, or whether these timeframes were set by the Synthesis research institute (in the method paper it says that the goals were defined *in coordination with the AMS* (Holl, et al., 2018, p. 4)). According to an AMS-internal definition, a person is regarded as *long-term unemployed* starting from a period of one year (AMS, 2019), so that this notion cannot be used to explain the defined timeframes of seven, respective 24 months.

A different definition of these two objectives with different timeframes for the observation period used (seven resp. 24 months) and/or for the employment objective (90 days resp. 180 days) would imply different coefficients in the implementation of the model and thus a different composition of the groups formed, since the calculated probability always depends on the corresponding definition of objectives. (For example, it is obvious that the probability of achieving employment for at least two months in the next seven months is greater than the probability for at least three months, since all those who achieve the three-month goal achieve the two-month goal a fortiori, and thus the success population is larger.)

Thus, these are acts of definition which are based on human decisions and could therefore have been different. The categorization of job-seekers who are algorithmically classified and the concrete composition of the groups both depend on these human decisions.

### 4.3 The Thresholds

A logistic regression model a priori does not yet determine a classification into different categories. It initially merely estimates the probabilities of achieving certain goals. The classification of individual job-seekers, and the partitioning of all job-seekers into different groups, is carried out via the introduction of cut-off points. (Hastie, et al., 2008) The threshold value for being categorized into Group A is $f_1 > 0{,}66$, the threshold value for Group C is $f_3 < 0{,}25$. Thus, the two probabilities are first calculated for the corresponding job-seekers, and then the classification in Group A, B or C is determined along these threshold values.

Other thresholds would therefore imply a different decomposition into the groups. If, for example, the threshold value for Group C is higher than 0,25 then, obviously, more people would automatically fall into this group. It is therefore crucial for the configuration of the three groups exactly where the threshold values are set.

In the method paper it is stated that the threshold values 0,25 and 0,66 were chosen *considering the accuracy of the model.* (Holl, et al., 2018, p. 6) This means that these values are intended to make the accuracy as good as possible, so that the rationale behind them is to be able to correctly classify as many people as possible from the available data from the past in retrospect.

In the method paper, there is no indication of the fact that one reason for the thresholds (and, thus, for the specific composition of the groups) could lie in the factual context of the job-seekers' labour market situation and their respective needs, so that it can be assumed that the issue here really is only one of accuracy and not a qualitative assessment of which measures (such as the BBEN, see below) would be suitable for which people.

In summary, the classification of job-seekers along the AMS Algorithm is based on three conceptual building blocks that are designed from specific perspectives and are therefore neither objective nor neutral. The data previously collected reflects to a certain extent the labour market situation in the past (see below), so that bias and unequal treatment on the labour market is naturally present and is inscribed in the model. Behind the definition of the two goals with regard to which the probabilities $f_1$ and $f_3$ are calculated, are specific values and very specific perspectives, which were not further discussed in the method paper. The thresholds are predefined cut-off points that shape and define the group affiliation of each job-seeker.

The algorithmic classification (and thus, the three building blocks) can have a significant impact on the situation job-seekers find themselves in. It is planned that being assigned to Group C will lead to removal from the AMS-internal resource system and a transfer to the external format of the BBEN (see below).

## 5 Group C

The method paper does not provide sufficient information to reasonably assess the implications of the categorization. The coefficients in Fig. 1 above, for example, that were published in the method paper, only apply to the base population of job-seekers who have an employment history that is statistically fully recorded, which means that there are no gaps in their data (see above). Furthermore, these coefficients apply to the short-term goal $f_1$ which only determines whether or not a person will be assigned to Group A. More interesting and important, however, would be the coefficients for the *long-term goal* $f_3$, since $f_3$ is used to determine whether a person is assigned to Group C. Nonetheless, some statistics in the method paper do give an idea about the degree of vulnerability of Group C.

### 5.1 Fragmented Employment History

Of the sub-population of cases of job-seekers with a *fragmented* employment history and residence in Vienna, the classification algorithm assigns a total of 29% to Group C. (Holl, et al., 2018, p. 15) It says in the method paper that the employment history of a job-seeker is *fragmented*, for example, when it comes to young people (having had no significant employment history), immigrants (having had no employment history in the Austrian job market), or people returning to work after a long period of time, (Holl, et al., 2018, p. 5) so that it may be assumed that women who do not have a continuous employment history due to childcare in the 48 months prior to the start of the respective AMS case belong to this statistical sub-population.

These people are excluded from the *base population* (see above): They form the counterpart to the norm of the continuously employed persons with social security status. This shows another dimension of inequality: It may be assumed that not being "datafied" enough in the past correlates to a high probability of being assigned to Group C. To fully assess this, more statistical data about the actual distribution of the job-seekers via the classification is needed.

### 5.2 Gendered Inequality

Initially, Johannes Kopf, who is in the Executive Board of the AMS, countered the accusations that this algorithmic system with its negative coefficients for the data entry *Gender: Female* is discriminatory by saying that there was a labour market policy goal in Austria anyway of spending 50% of the AMS support resources on women. (Wimmer, 2018) Interestingly enough, the counter-argument is not that the algorithmic classification and its consequences *do not discriminate* against women – instead, he claimed that women were the focus of a 50% support policy quota that interferes in that sense with the redistribution of the AMS resources via the algorithmic classification. However, with the new labour market policy goals set by the Austrian government, this policy goal has now been withdrawn. (Szigetvari, 2019)

Concrete and extensive statistics on gender proportions in Groups A, B and C were not published. Kopf's statement that there are "disproportionately many" women in the most eligible Group B (Kopf, 2018b) does not provide any information on the actual impact of the gender dimension. Model variants implemented once with and once without gender as a relevant feature would have to be compared, and these variants do not exist. Also, since gender inequality on the labour market was statistically found to have a negative influence on women's job placement (reflected by the negative coefficient of the feature *Gender: Female*), a predictive model that does not consider gender would therefore not be as accurate – which points to the dilemma that a high degree of accuracy in predictions that are based on data from the past just translates to the model being as good at discriminating structurally as the labour market is, see below.

The only gendered statistic that can be found in the method paper is the following: Of the fully datafied base population of cases of women, estimated at the beginning of the case at the AMS, 5% are categorized in Group C - which is more than one and a half times the size of the corresponding sub-population of men. (Holl, et al., 2018, p. 15)

### 5.3 The BBEN

In order to assess the impact of the algorithmic classification of job-seekers it is crucial to identify what exactly happens to those job-seekers who are assigned to Group C. As mentioned above, the categorization of job-seekers into three groups is supposed to be a step on the way to categorize and reorganize the distribution the AMS support resources. Those job-seekers with low chances on the labour market should get access to *different resources* (Kopf, 2018a) that are not provided by the AMS itself. According to § 32 AMSG and § 34 AMSG, the AMS can cooperate with external (non-AMS) institutions (BBE – *Betreuungs- und Beratungseinrichtungen*) when it comes to support services that the AMS cannot provide, the reason being that offering these services would be *unsuitable* or *uneconomical*. (Weber, et al., 2019, p. 5) One such external service format mentioned earlier, the BBEN (*Beratungs- und Betreuungseinrichtung Neu*), was tested in a pilot project in late 2017 and in 2018 in several regions in Austria, and subsequently evaluated externally by a consultancy agency. This document is available online on the AMS research platform *Forschungsnetzwerk*.

It says in the evaluation that the background for the BBEN is the planned segmentation of job-seekers. (Weber, et al., 2019, p. 26) As the AMS will focus its own resources on people with *medium* chances on the labour market, this leads to a need for a new support format for people with especially low chances on the labour market. (Weber, et al., 2019, p. 23) The BBEN's target group are has job-seekers with *multiple job placement difficulties*, and its constitutive purpose is to *preserve the chances of job placement*. (Weber, et al., 2019, p. 15)

In late 2017 and throughout 2018, the BBEN were tested for a specific subgroup of job-seekers (that were not algorithmically classified, as the algorithmic classification was only introduced in 2019): People with a 2-year long unemployment case history at the AMS who fulfill at least two of the following criteria - being at an age above 45, having a low degree of education, and having a disability. (Weber, et al., 2019, p. 15) It says explicitly in the document that the BBEN will be extended to people who have a *probability of less than 25% of achieving job placement for six months within the next 24 months*, (Weber, et al., 2019, p. 16) which is precisely the defining condition $f_3 < 0{,}25$ for Group C. Therefore, it can be concluded that being assigned to Group C will lead to being removed from access to the internal AMS support resources to the external BBEN resources.

The evaluation assessed, among other things, the satisfaction on the part of the AMS and selected participants of the BBEN. Job-seekers who had participated on one or more offers of a BBEN for six months or longer were interviewed and the result of the evaluation is said to be consistently positive: 83% of the surveyed participants were *very content* with the BBEN. (Weber, et al., 2019, p. 9) However, one characteristic of this external format is voluntary participation: After a single mandatory information event at the beginning, the assigned candidates only participate on a voluntary basis. (Weber, et al., 2019) Therefore, the composition of the surveyed group of job-seekers should be reflected accordingly: If only those candidates are interviewed who have *voluntarily* used the BBEN resources for at least six months, they are naturally satisfied with the BBEN resources, as otherwise they would not have used them. Of 5700 candidates of the pilot project in 2017 and 2018, just under 47% (2675 candidates) took advantage of the BEEN resources for at least six months. (Weber, et al., 2019, p. 16;20)

Furthermore, it is noteworthy that the most frequently cited reason for satisfaction by the surveyed participants is the possibility of one-on-one meetings within the BBEN (Weber, et al., 2019, p. 10), so that the assumption that a one-on-one meeting with somewhat more time capacities within the AMS system could also lead to the same degree of satisfaction is not far off. According to the evaluation, 98% of the interviewed participants took advantage of individual meetings and counselling. (Weber, et al., 2019, p. 10)

The planned segmentation of job-seekers, which is to be introduced, among other things, in order to save internal counselling resources of the AMS and, therefore, in order to reduce the case workers' workload, also encounters a certain degree of reluctance within the system. On the one hand, it was pointed out in the evaluation that the BBEN format leads to a clear relief of the AMS case workers and provides remedy against the pressure to reduce the duration of the AMS cases. (Weber, et al., 2019, p. 34) It can therefore be concluded that admitting a job-seeker to the BBEN leads to the internal ending of the respective case, so that the case workers' success rates in ending cases turn out to be increased. Thus, AMS case workers in fact only have advantages with the introduction of the BBEN. Accordingly, 86% of the interviewed case workers consider it a relief to have this new external resource. (Weber, et al., 2019, p. 71)

Yet on the other hand, a total of 37% believe that the group of job-seekers that are assigned to the BBEN should rather remain within the internal AMS system. The evaluation speaks of "only 37%" (Weber, et al., 2019, p. 71) – but in view of the exclusive advantages of the BBEN for the AMS case workers, this number can also be regarded as very high: Almost 40% of the interviewed case workers are therefore in favour of keeping the BBEN-group within the AMS instead, despite the fact that this externalization clearly relieves them of some of their workload.

It is noted at one point in the evaluation that caution is advised at the transition of BBEN-participants returning to "intensified AMS" counselling to avoid the BBEN becoming a "one-way street". (Weber, et al., 2019, p. 73) This hint is to be read in view of the consistently very affirmative tone and optimistic outlook of the evaluation. While the positive aspects are repeatedly elaborated, the evaluation clearly sees a potential problem here. Being assigned to a BBEN, which, as it is planned, is equivalent to being assigned to Group C, could potentially be a "one-way street".

# 6 Intersectional Inequality

Intersectional discrimination occurs when a person experiences discrimination on the basis of several interrelated and compounding dimensions of inequality. (Holzleithner, 2016) (Crenshaw, 1989) (Uccellari, 2008) The algorithm that estimates the probability of achieving the two job placement goals $f_1$ and $f_3$ has been designed in terms of AMS data from the past. In retrospect, it was thus possible to determine *which groups* of people with *which features* (that are, and therefore, *can be* recorded in data) managed to achieve job placement *when* and *for how long*. (All these questions are encoded in the definition of $f_1$ and $f_3$, see above.) The negative coefficients of certain data entries, such as *Gender: Female*, age above 30, disability, childcare responsibilities, non-EU citizenship, show that these factors, under the underlying model assumptions, have had a negative impact on achieving job placement of the people recorded. Although only the concrete coefficients for $f_1$ for one sub-population were published, it can be assumed that the negativity of these coefficients can generally be found in all model variants. (Other coefficients for other model variants could not be found, and a request to Synthesis for even example coefficients for $f_3$ was rejected.)

The underlying assumption behind the development of a model using logistic regression is that the target probability can be estimated sufficiently well by the variables (features) included. (Hastie, et al., 2008) Each coefficient is to be interpreted in the sense that having the corresponding data entry (e.g. *Gender: Female*) viewed on its own (with all other features fixed) has the corresponding influence on the probability (described by the coefficient), in the example above -0,14. (Hastie, et al., 2008)

The statistical finding that the feature *Gender: Female* has a negative coefficient thus shows that there is a *structural* disadvantage in the labour market: Two job-seekers with otherwise completely identical data entries have statistically different success rates with regard to job placement. The feature *Gender: Female* with otherwise unchanged data has a negative effect.

This knowledge could potentially open up an emancipatory moment in the use of such predictive technologies. Indeed, the predictive model does not calculate the individual *chances* a person *has* on the labour market, as from the very beginning these are collective groups formed along datafied (and datafiable) categories. *Rather, the model shows in the form of the negative coefficients how and to what extent the Austrian labour market differentiates and discriminates structurally on the basis of various features.* Precisely this structural dimension of disadvantage, which has nothing to do with the specific individual, is embedded in the statistical significance of the different included features.

The model therefore does not calculate the individual *chances* that a person has, but *reflects* the structural situation on the labour market with which this person will be confronted when searching for a job. As an analysis of the Austrian labour market and its discriminatory tendencies, this model with its coefficients could thus be an insightful tool for distributing support resources using a bottom-up approach, see below. The current use of the model does the opposite, however, in that individuals are subjected to the collective disadvantage of their non-voluntary membership to a group formed via data categories that is discriminated against structurally.

- In a first step, the *AMS Algorithm* depicts intersectional discrimination on grounds of gender, parenthood, age, citizenship and disability that occurs in the job market via the negative coefficients in the predictive model.
- In a second step, the individual facets reinforce each other to create yet another dimension of disadvantage: Being a person at the intersection of different axes of vulnerability can lead to being assigned to the less eligible Group C.

If job-seekers are exposed to structural discrimination on the labour market to a certain (predefined) extent, namely $f_1 < 0{,}66$ while at the same time $f_3 > 0{,}25$, then they fall into the Group B and have access to all AMS support resources. If, according to the model, their disadvantages exceed the predefined threshold value, so if $f_3 < 0{,}25$, for example by cumulation of various *personal features* with negative coefficients, then they fall into Group C, which was defined as less eligible within the AMS resources system according to the "efficiency" criterion.

The fact that people are subjected to discrimination because of their age, gender, care responsibilities, nationality (and thus implicitly because of racism in the labour market), is observed, then confirmed in the statistical model and finally reinforced by means of the classification and the corresponding efficient distribution of AMS resources.

## 7 Merely a Tool? – Individualizing the Problem

It is emphasised by the AMS that the model is merely a tool for decision-making and is not formally binding, so that the *individual final decision* about the classification should remain with the (human) case worker. (Kopf, 2018a) However, research on the use of automated decision-making tools has long shown that these tools are often neither well-understood nor questioned by users. (Parasuraman & Riley, 1997)

This referral on the part of the AMS to the *individual final decision* of the respective case worker is essential insofar as the planned use of the predictive model in this way escapes the corresponding legal conflicts on equal treatment and anti-discrimination. The AMS is subject to the Equal Treatment Law (GlBG: *Gleichbehandlungsgesetz*), which prohibits unjust unequal treatment based on gender, parenthood, ethnic background, age and several other protected features. Unequal treatment because of higher or lower education, for example, does not fall under the legal definition of discrimination. An unlawful unequal treatment occurs if a less favourable treatment takes place solely because of one or more of the aforementioned *protected* features. These features, however, are *explicitly implemented* in the algorithm via their corresponding coefficients, so that being assigned to a less eligible group due to an intersectional cumulation of *negative* coefficients might be considered to be exactly such an unequal treatment. A normative and formally binding acceptance of the algorithmic classification without referring to the individual final decision could therefore be assumed to not be compatible with this law.

The AMS bypasses this legal conundrum by pointing to the individual case workers. So, if a discriminatory use of this technology can be proven, the case workers will be held responsible, since, ultimately, they are said to make the final decision. It will therefore be required of the case workers that they always make the right decision as to whether or not they follow the algorithmic classification, all that in addition to their increased workload (with which, among other factors, this algorithm was justified). *The issue of structural unequal treatment, which is first reflected by and then inscribed into the model, is being argumentatively reduced to the individual level of the case workers*.

Thus, there is a certain field of tension, so that on the one hand this model was obviously developed in order to be extensively used in practice (otherwise the almost 240.000 Euros spent (Kopf, 2018) would not be justifiable in times of efficiency increase), and on the other hand it is always emphasized that the use will be restricted by the individual case workers to a non-discriminatory level of usage.

The model was designed in order to be able to provide a more *objective* (Kopf, 2018) assessment of job-seekers with *highly complex mathematical models* (Kopf, 2018c) and with more information (data) than the case workers on an individual level could ever have, (Kopf, 2018b) and yet the case workers are expected to have some sort of meta-intelligence to be able to judge whether or not to use the model in specific situations with specific job-seekers.

## 8 Efficiency

As mentioned above, the rationale for using the predictive model is an accompanying increase in *efficiency*. This refers to an efficiency on two intertwined levels: At a macro level, the overarching objective of this labour market policy measure is job placement for as many people as possible. (Kopf, 2018a) Furthermore, at the micro-level of the operational processes within the AMS, case work is to be transformed to the extent that costly resources, such as one-to-one counselling, can be focused more strongly on Group B. Johannes Kopf speaks of *reducing contact intensity of this group.* (Kopf, 2018a) Thus, valuable resources, such as the above-mentioned counselling resources should be used where they are most *efficient* in terms of maximizing the number of job placements according to the defined *short-term goal* and *long-term goal*.

This labour market policy objective, i.e. job placement for as many people as possible, is to be understood as a *defined* objective. A different conception of objective therefore would yield a different conception of efficiency in relation to this objective. According to a bottom-up strategy, for example, the most important objective of labour market policies could be to support those job-seekers who, for a variety of reasons, have particularly low chances on the labour market. (Crenshaw, 1991) The people of Group C who are currently being handled as almost negligible collateral damage in the current use of the model would thus become the starting point and centre of the AMS labour market policy efforts.

According to Judith Pühringer, Executive Director at *arbeit plus*, a network of over 200 non-profit, labour market orientated Social Integration Enterprises in Austria, efforts along this same bottom-up approach were at least to some extent the focus of the AMS operations until the introduction of the algorithmic classification. In October 2018, when the AMS Algorithm was widely discussed in the media, she said: "*Currently people who have the greatest need receive the most support. Now we are moving away from this logic.* […] *The* [new] *focus is on the middle segment*". (Szigetvari, 2018)

The concept of *efficiency*, which goes hand in hand with the introduction of the categorization of job-seekers via the AMS Algorithm, thus becomes discriminatory in its impact only as a result of the unquestioned fact that the particularly disadvantaged people of Group C can be regarded as negligible with regard to the internal AMS resources, so that they automatically fall outside the labour market objectives, and thus all the more outside the concept of *efficiency*.

## Conclusion

This paper has elaborated that the AMS Algorithm is based on three building blocks, each with a conceptual and a concretely implemented dimension: The past to which one refers when making predictions is abstracted and reflected in the data that is available. The outlook on the future that is used to derive measures is found in the target variables, i.e. in the two goals. The cut-off points reflect a valuation along which it is decided who shall belong to which group. These three elements are based on human decisions and are therefore neither objective nor neutral.

The field of STS is well aware that the production of scientific knowledge is a social undertaking that can thus be examined through a social, as well as a political lens, so that the situatedness of the scientists can, and should, always be taken into account. (Haraway, 1988) In the case of the AMS Algorithm, behind the proclamation of these *highly complex mathematical techniques*, which is in line with the currently prevalent Big Data hype (boyd & Crawford, 2012) there are quite clearly visible valuations, decisions and presumptions, as shown above. As intensified data collection and data analysis are to be expected in the future (as well as in the present), and as phenomena are being transformed and reduced to datafiable quantities, (Kitchin, 2014) which has been shown to be much more than just reduction, but a process that impacts ontological dimensions, (Mol, 2002) the epistemological foundations of Big Data guided methods are to be critically examined. (Prietl, 2019) The complex problems that can occur when working with large amounts of data (Busch, 2014), such as missing data or data gaps, (Harrell, 2015) can result in bias that may be prevalent in the data itself, for data cannot be objective, and is always “cooked” (Gitelman, 2013) as well as in implemented algorithms. (Friedman & Nissenbaum, 1996)

The AMS Algorithm is a concrete and therefore an illustrative and instructive example of how and where to locate presumptions in algorithmic techniques. As one characteristic of such technologies is scale, (O'Neil, 2016) meaning that biased tendencies can be easily and efficiently transferred to large numbers of people, inequalities can be reinforced and amplified on a large scale. (Eubanks, 2018) If the algorithmic technique is based on a predictive model (Hofman, et al., 2017) so that socio-political governance measures (Rieder & Simon, 2016) are derived from the prediction, (Jasanoff, 2005) the prediction itself can be highly biased, (Angwin, et al., 2016) and the derived measures can reinforce inequalities as a (literal) self-fulfilling prophecy, or feedback loop, (Ensign, et al., 2018) especially when imposed on vulnerable target groups.

## References


AMS, 2019. *Arbeitsmarktdaten und Arbeitsmarktforschung: Fachbegriffe.* [Online] Available at: https://www.ams.at/arbeitsmarktdaten-und-medien/arbeitsmarkt-daten-und-arbeitsmarkt-forschung/fachbegriffe#langzeitbeschftigungslose [Retrieved 3 5 2019].

Angwin, J., Larson, J., Mattu, S. & Kirchner, L., 2016. *Machine Bias. There's software used accross the country to predict future criminals. And it's biased against blacks..* [Online]
Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
[Retrieved 2 5 2019].

boyd, d. & Crawford, K., 2012. CRITICAL QUESTIONS FOR BIG DATA. *Information, Communication & Society*, Issue 5, pp. 662-679.

Busch, L., 2014. A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large-Scale Data Sets. *International Journal of Communication*, pp. 1727-1744.

Crenshaw, K., 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, pp. 139-167.

Crenshaw, K., 1991. Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color. *Stanford Law Review*, July, pp. 93-118.

Ensign, D. et al., 2018. Runaway Feedback Loops in Predictive Policing. *Proceedings of Machine Learning Research*, pp. 1-12.

Eubanks, V., 2018. *Automating Inequality: How high-tech tools profile, polica, and punish the poor.* New York: St. Martin's Press.

Friedman, B. & Nissenbaum, H., 1996. Bias in Computer Systems. *ACM Transactions on Information Systems (TOIS)*, Issue 3, pp. 330-347.

Gitelman, L., Hrsg., 2013. *"Raw Data" Is an Oxymoron.* Cambridge: The MIT Press.

Goodfellow, I., Bengio, Y. & Courville, A., 2016. *Deep Learning.* s.l.:MIT Press.

Haraway, D., 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, Issue 3, pp. 575-599.

Harrell, F. E. J., 2015. *Regression Modeling Strategies: With Application to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* 2 Hrsg. s.l.:Springer.

Hastie, T., Tibshirani, R. & Friedman, J., 2008. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2 Hrsg. s.l.:Springer.

Hofman, J., Sharma, J. & Watts, D., 2017. Prediction and explanation in social systems. *Sciencs*, 3 2, pp. 486-488.

Holl, J., Kernbeiß, G. & Wagner-Pinter, M., 2018. *Das AMS-Arbeitsmarktchancen-Modell. Dokumentation zur Methode.* [Online]
Available at:
http://www.forschungsnetzwerk.at/downloadpub/arbeitsmarktchancen_methode_%20dokumentation.pdf
[Retrieved 4 3 2019].

Holzleithner, E., 2016. EU-rechtliche Bestimmungen zum Diskriminierungsverbot: Grundlagen und Anwendung. In: A. Scherr & A. El-Mafaalani, Hrsg. *Handbuch Diskriminierung.* Wiesbaden: Springer.

Hong, R., 2016. Soft skills and hard numbers: Gender discourse in human resources. *Big Data & Society*.

Jasanoff, S., 2005. Technologies of humility. In: C. Mitcham, Hrsg. *Encyclopedia of Science, Technology and Ethics.* New York: Macmillan Reference, pp. xix-xxvii.

Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. *Big Data & Society*.

Kitchin, R., 2017. Thinking critically about and researching algorithms. *Information, Communication & Society*, pp. 14-29.

Kopf, J., 2018a. *Der Standard. Interview and Article by András Szigetvari: AMS-Vorstand Kopf: "Was die EDV gar nicht abbilden kann, ist die Motivation".* [Online]
Available at: https://derstandard.at/2000089096795/AMS-Vorstand-Kopf-Menschliche-Komponente-wird-entscheidend-bleiben?ref=article
[Retrieved 17 1 2019].

Kopf, J., 2018b. *JohannesKopf.at: Wie Ansicht zur Einsicht werden könnte.* [Online]
Available at: https://www.johanneskopf.at/2018/11/14/wie-ansicht-zur-einsicht-

werden-koennte/
[Retrieved 06 2019].

Kopf, J., 2018c. *Der Standard. Kommentare der anderen: AMS-Algorithmus - von der Ansicht zur Einsicht.* [Online]
Available at: https://derstandard.at/2000091365385/AMS-Algorithmus-von-der-Ansicht-zur-Einsicht
[Retrieved 2 4 2019].

Kopf, J., 2018. *Futurezone. Interview and Article by Barbara Wimmer: AMS-Chef: "Mitarbeiter schätzen Jobchancen pessimistischer ein als der Algorithmus".* [Online]
Available at: https://futurezone.at/netzpolitik/ams-chef-mitarbeiter-schaetzen-jobchancen-pessimistischer-ein-als-der-algorithmus/400143839
[Retrieved 17 January 2019].

Latour, B., 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory.* Oxford: Oxford University Press.

Mol, A., 2002. *The Body Multiple: Ontology in Medical Practice.* Durham: Duke University Press.

O'Neil, C., 2016. *Weapons of Math Destruction; How Big Data Increases Inequality and Threatens Democracy.* New York: Crown Publishing Group.

Parasuraman, R. & Riley, V., 1997. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Issue 39, pp. 230-253.

Prietl, B., 2019. BIG DATA: Inequality by Design?. *Proceedings of the 2nd Weizenbaum Conference on Digital Inequality*.

Rieder, G. & Simon, J., 2016. Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society*, Issue 1.

Szigetvari, A., 2018. *Der Standard: AMS bewertet Arbeitslose künftig per Algorithmus.* [Online]
Available at: https://derstandard.at/2000089095393/AMS-bewertet-Arbeitslose-kuenftig-per-Algorithmus
[Retrieved 9 5 2019].

Szigetvari, A., 2019. *Der Standard: Sozialministerium streicht 50-Prozent-Förderziel des AMS für Frauen.* [Online]
Available at: https://derstandard.at/2000098639725/Sozialministerium-strich-50-Prozent-Foerderziel-fuer-Frauen-am-Jobmarkt
[Retrieved 14 3 2019].

Uccellari, P., 2008. Multiple Discrimination: How Law can Reflect Reality. *The Equal Rights Review*, pp. 24-49.

Weber, F., Hager, I., Krüse, T. & Reidl, C., 2019. *Evaluierung des*

*Betreuungsformates für Personen mit multiplen Vermittlungshindernissen (BBEN).* [Online]
Available at: http://www.forschungsnetzwerk.at/downloadpub/2019_BBEN_BBEN-ams_final.pdf
[Retrieved 2 5 2019].

Wimmer, B., 2018. *Futurezone: Der AMS-Algorithmus ist ein "Paradebeispiel für Diskriminierung".* [Online]
Available at: https://futurezone.at/netzpolitik/der-ams-algorithmus-ist-ein-paradebeispiel-fuer-diskriminierung/400147421
Retrieved 8 2 2019].

Zheng, A. & Casari, A., 2018. *Feature Engineering for Machine Learning. Principles and Techniques for Data Scientists.* s.l.:O'Reilly.