

A Scandal on AI in Administration, Again

Sarah de Heer

2023-07-25T00:42:32

After the infamous [Dutch benefits scandal](#) – that started in the 2000s and for which compensation of the wronged parties has still to be achieved –, the Netherlands are yet again the scene of wrongful application of an algorithm by the government. This time, the main actor is the *Dienst Uitvoering Onderwijs* (**DUO**), the Dutch agency responsible for the allocation and payment of student loans to those enrolled in Dutch higher education. Specifically, DUO used an algorithm in their enforcement task, namely to verify whether the student loans have been rightfully allocated. In 2012, DUO commenced the use of this ‘in-house’ algorithm, which the Minister of Education – under whose responsibility DUO falls – [halted](#) on 23 June. The developments in the Netherlands epitomize the promises and pitfalls of further integrating automated decision-making (**ADM**) into public administration. On the one hand, ADM – sometimes labelled ‘artificial intelligence’ – is cheap and promises efficiency gains. On the other hand, ADM systems may be error-prone when facing the complex realities of societal life and legal ambiguity.

This contribution explains the facts of the use of the algorithm used by DUO, which is largely based on news articles ([here](#) and [here](#)). Subsequently, this blogpost outlines the peculiarities of the algorithm, as currently known, and discusses which fundamental rights and principles are at stake. Last, this piece proposes a research agenda that recommends the approaches ahead to ensure adequate use of AI by public administration. Thus, this Section zooms out and focusses on the broader picture, namely the use of automated decision-making systems – which are AI-based tools that uses profiling to predict an outcome – by public administration. This contribution does not hold that the use of such algorithms should be ceased but rather argues that certain safeguards should be put in place to guarantee fundamental rights and principles. Specifically, this contribution calls for i) public administrators working with ADM systems to have sufficient knowledge on how these systems work, ii) preventing tunnel vision, and iii) developing transparent ADM systems to obtain accurate and representative outcomes.

Setting the scene – DUO’s use of an ADM system in administrative decisions

Students enrolled in Dutch higher education may be entitled to receive student loans – and depending on the income of the student’s parents or legal guardians a grant conditional upon successfully obtaining the degree – to finance their studies. The amount of the student loans varies and depends on whether the student decides to live with their parents or legal guardians or rather elsewhere. Concretely, students are entitled to a larger sum, provided i) they are formally registered on a different

address than that of their parents or legal guardians, and ii) they are – in fact – residing at the registered address. To verify whether these student loans were rightfully paid, DUO made use of an ADM system to predict which students most likely falsely claimed to live away from home – but effectively opt to live with their parents or legal guardians –, and thus have falsely received the larger amount of student loans. This blogpost divides the procedure to detect alleged fraudsters four steps.

First, the ADM system identifies through means of profiling possible fraudulent students, which amounts to 10,000s potential files (*pre-selection phase*). Second, five DUO employees sift through these files to verify i) the living surface of the property, ii) how many people are registered on the address, and iii) whether the student lives with a family member. This assessment results in reducing the number of files to 1,000 (*selection phase*). Third, these shortlisted 1,000 cases are to expect a home visit by DUO inspectors, who determine whether the student effectively lives on the registered address. During such a house visit, which may not take up much time, the inspectors assess daily and sometimes seemingly trivial affairs to reach their conclusion. In particular, the inspectors may open wardrobes to ascertain whether there are enough clothing items, take note of the amount of coursebooks, whether the bed has been slept in and count the toothbrushes in the toothbrush holder. Unfortunately, these house visits are also conducted during holidays and during daytime – when students are normally not home. Thus, after visiting the registered address three times – in vain – to gain access to the premises, the inspectors proceed with canvassing the student's neighbours. Concretely, the inspectors may ask the neighbours whether they are aware of any students living in the neighbourhood. After the home visit or the house-to-houses, the inspectors draft a report with their findings as to whether a student effectively resides at the registered address (*enforcement phase*). Fourth and last, DUO issues a decision based on the inspectors' report. Where DUO finds that the student does not reside at the registered address, and thereby illegally received larger sums, the student is required to not only repay the excess amount but also a fine (*administrative phase*). Receiving such a decision has substantially detrimental effects on students – who are generally restricted to a limited budget –, as the repayments and the fine may amount to thousands of euros and even upwards of 10,000 euros, which is thus a hefty amount on a student budget. Since 2012 – the year in which DUO started to use their own algorithm –, DUO has discovered 9,923 cases of probable fraud by using this method.

Peculiarities stemming from the used ADM system – which fundamental rights and principles at stake?

While the use of ADM systems for administrative decisions is a trend observed around the globe – and thus not a surprising development in the Netherlands –, this blogpost identifies three worrisome characteristics regarding the individual's enjoyment of fundamental rights and principles.

First, the above procedure may jeopardise the *right to good administration*. Where the inspectors only conducted the house-to-house, the report would be solely founded on the canvass performed, which can hardly be seen as conclusive evidence. An example of a statement is:

'You ask whether a student is living here. No. I only encounter these people and their child. Apart from them, I don't see anybody. That is all I know.' (author's own translation)

Surprisingly, DUO would deem two of such statements sufficient to reach the conclusion that student did not live at the registered address. At the eve of the reintroduction of basic grants that replaces student loans, it is likely that DUO will intensify its reinforcement task, which makes this concern even more pertinent. Further, the administrative appeal procedure at DUO shows the tendency to rely on the neighbours' statements obtained during the house-to-house irrespective of the counterevidence submitted by the student, which demonstrates DUO's inclination to rely unequivocally on their inspectors' findings. This illustrates risks to the duty to state reasons that requires DUO to clarify how they reached the decision that student does not effectively reside at the registered address. Moreover, DUO adopts such a decision without having a dialogue with the concerning student, which leaves the right to be heard imperilled. Additionally, since these students did not obtain a reasoned decision explaining why they are deemed to be a fraudster, their *right to an effective remedy* is thwarted.

Second, moving on to the specific features of the ADM system, DUO has devised the algorithm underlying the ADM system, which enables the profiling of students receiving student loans. Specifically, DUO fed the algorithm so-called 'risk indicators', which included the student's age, the student's level of education, the student's address, the address of the student's parents or legal guardians, and the address of the student's educational institution. It is likely that when a student matches one or more of these 'risk indicators', it increases the student's likelihood of being deemed to have committed fraud with the student loans. However, it remains unclear what these 'risk indicators' precisely entail, which poses a threat to the *principle of legal certainty* and the *principle of legitimate expectations*. Moreover – and potentially more disturbingly since it exacerbates the above –, even though this algorithm was not based on any legislative measure but rather on DUO's own initiative, the algorithm created by DUO itself, nevertheless, formed the foundation of the harmful decision to repay the surplus amounts and an additional fine. To sum up the above, due to these unknown risk indicators, the ADM system itself brings serious perils to the overall *principle of transparency*.

Third, zooming in on the risk indicators, DUO seems to have actively prevented direct discrimination, as it did not feed the algorithm data on nationality or on the country of origin. However, DUO did make use of proxy data that may reveal such sensitive data, which – in its turn – may result in indirect discrimination. Thus, *the right to non-discrimination* may be hampered. This is also evidenced by the sampling performed by NOS op 3, a Dutch broadcasting programme, and Investico, a Dutch platform of investigative journalism. They approached 70 lawyers representing students who have received a decision from DUO stating that they did not reside

at the registered address, and who thereby were confronted with the hefty sums, out of which 32 lawyers shared 376 files. Remarkably, 367 decisions were directed to students with a so-called ‘migration background’, which amounts to 97% – a noteworthy overrepresentation. As a result of the outcome of the research performed by NOS op 3 and Investico, the Dutch Data Protection Authority has [initiated investigations](#) as regards the processing of personal data for the creation of the algorithm.

Conclusion – how to fortify fundamental rights against AI?

Taking a step back to appreciate the general trend, namely the increased use of ADM systems by public administration aimed at partly or fully replacing administrative decisions that may have significantly adverse effects on the individual, the common denominator is formed by the government pursuing a noble aim – be it fraud detection, social benefits allocation, or crime prevention – but at the same time lacking the required expertise to adequately use these novel technologies in a manner that respects citizens’ fundamental rights and principles. This means that there is a glaring need to research how to ensure fundamental rights and principles seeing the apparent risks posed by the algorithms used.

Thus, I argue, *first*, that the public administrators who use these AI-based tools should have sufficient knowledge on how the algorithm works and the risks and benefits thereof. To this end, the AI Act, as amended by [European Parliament](#), could be helpful, as it requires AI literacy for those working with AI-based systems. Specifically, the European Parliament holds that such a sufficient knowledge could be achieved by providing training on, amongst others, basic notions, and the functioning of the AI-based device. *Second*, tunnel vision should actively be discouraged. In this regard, the legislator may play a pivotal role by, for example, prohibiting the mere reference to the (brief) findings of public administration that confirm the outcome of the algorithm when presented with sufficiently substantiated counterevidence. *Third*, the initiative to create and use ADM systems should stem from legislative measures – as opposed to administrative actions –, especially when these ADM systems may negatively affect individuals. Specifically, as regards the data used to develop the algorithm, they should be accurate to achieve representative outcomes, be construed in a transparent manner, and stem from a sufficient sample size. For example, the legislator may demand the publication of the metadata of the data used to create the ADM system. Not only will the developer of the ADM system and external parties become aware of the data used, but – and perhaps more importantly – the metadata may also demonstrate the data missing, which may provide insights as regards the accuracy and representation of the outcomes of the ADM system. This holds even more true when using data that enables profiling, namely the ‘risk indicators’.

