



# XENOPHOBIC MACHINES

DISCRIMINATION THROUGH UNREGULATED USE OF ALGORITHMS IN THE DUTCH CHILDCARE BENEFITS SCANDAL

**Amnesty International is a movement of 10 million people which mobilizes the humanity in everyone and campaigns for change so we can all enjoy our human rights. Our vision is of a world where those in power keep their promises, respect international law and are held to account. We are independent of any government, political ideology, economic interest or religion and are funded mainly by our membership and individual donations. We believe that acting in solidarity and compassion with people everywhere can change our societies for the better.**



Cover illustration: Tosca van der Weerden

© Amnesty International 2021

Except where otherwise noted, content in this document is licensed under a Creative Commons (attribution, non-commercial, no derivatives, international 4.0) licence.

<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

For more information please visit the permissions page on our website: [www.amnesty.org](http://www.amnesty.org).

Where material is attributed to a copyright owner other than Amnesty International this material is not subject to the Creative Commons licence.

First published in 2021 by Amnesty International Ltd

Peter Benenson House, 1 Easton Street

London WC1X 0DW, UK

INDEX: EUR 35/4686/2021

OCTOBER 2021

ORIGINAL LANGUAGE: ENGLISH

**[amnesty.org](http://amnesty.org)**

**AMNESTY  
INTERNATIONAL**



# TABLE OF CONTENTS

<b>1</b>	<b>GLOSSARY</b>	4
<b>2</b>	<b>EXECUTIVE SUMMARY</b>	5
<b>3</b>	<b>SCOPE AND METHODOLOGY</b>	8
<b>4</b>	<b>BACKGROUND</b>	10
<b>5</b>	<b>DESIGN OF THE ALGORITHMIC DECISION-MAKING SYSTEM</b>	15
<b>6</b>	<b>DISCRIMINATION BY DESIGN</b>	18
6.1	LEGAL FRAMEWORK FOR NON-DISCRIMINATION AND RACIAL PROFILING	18
6.2	LEGAL ASSESSMENT OF THE ALGORITHMIC DECISION-MAKING SYSTEM	21
<b>7</b>	<b>FAILURE TO PROTECT HUMAN RIGHTS</b>	24
7.1	SOUNDING LIKE A BROKEN RECORD: PRIOR WARNINGS FOR HUMAN RIGHTS RISKS	24
7.2	OUT OF CONTROL: BLACK BOX SYSTEMS AND SELF-LEARNING ALGORITHMS IN THE PUBLIC SECTOR	26
7.3	KEPT IN THE DARK: A LACK OF TRANSPARENCY	27
7.4	HUMAN RIGHTS ON THE BACK BURNER: NO ASSESSMENT OF THE HUMAN RIGHTS IMPACTS	31
7.5	FAILING SUPERVISION: NO HUMAN RIGHTS OVERSIGHT	32
7.6	DISCRIMINATION IN PLAIN SIGHT: NO BAN ON RACIAL PROFILING IN THE NETHERLANDS	35
7.7	KAFKA'S CASTLE: A LACK OF EFFECTIVE REMEDIES AND REDRESS	37
<b>8</b>	<b>RECOMMENDATIONS</b>	39

# 1. GLOSSARY

**ALGORITHMIC SYSTEM** A system which uses a set of mathematical instructions or rules that calculate an answer to a problem or question.

**ALGORITHMIC DECISION-MAKING SYSTEM** An algorithmic system that is used in (support of) various steps of decision-making processes.

**AUTOMATED DECISION-MAKING SYSTEM** An algorithmic decision-making system where no human is involved in the decision-making process. The decision is taken solely by the system.

**SEMI-AUTOMATED DECISION-MAKING SYSTEM** An algorithmic decision-making system where a human is involved in the decision-making process. Often these systems are used to select cases for human review or to assist humans in the decision-making process by providing information and/or suggested outcomes.

**BLACK BOX SYSTEM** An algorithmic system where the inputs and outputs can be viewed, but the internal workings are unknown.

**PROXY** In computer programming, a proxy is a feature that is correlated to something else. In algorithmic systems, a seemingly neutral feature (such as postal code) may be correlated with a protected characteristic (nationality or ethnicity).

**RISK CLASSIFICATION MODEL** An algorithmic decision-making system used by the Dutch tax authorities to score applications of childcare benefits according to the risk of inaccuracy.

**SELF-LEARNING ALGORITHMS** Self-learning algorithms give algorithmic systems the ability to independently and autonomously learn from experiences over time, and to make changes to how they work without these changes being explicitly programmed by humans.

# 2. EXECUTIVE SUMMARY

Around the world, algorithmic decision-making systems are increasingly deployed to automate governmental tasks and to identify, detect, and punish people that meet predetermined profiles of alleged perpetrators and criminals. Many of these systems have been designed with a particular emphasis to combat fraud. In the Netherlands, fraud in social security became a prominent theme in political debate during the 1980s and 1990s. Consequently, social security enforcement policies and administrative sanctions were intensified. The childcare benefits scheme was introduced in 2005. This scheme is enforced by the Dutch tax authorities, who subjected it to these harsh social security law enforcement policies and the administrative sanction system. In 2013, an algorithmic decision-making system for fraud detection was adopted: the “risk classification model”. This algorithmic decision-making system included self-learning elements to create risk profiles of childcare benefits applicants who were supposedly more likely to submit inaccurate applications and renewals and potentially commit fraud. Parents and caregivers who were selected by this system had their benefits suspended and were subjected to investigation. The tax authorities had to prove the efficiency of their data-driven fraud detection methods, including the algorithmic decision-making system, by seizing enough funds from alleged fraudsters to cover the costs of the operation. This challenge created a perverse incentive to seize as many funds as possible, regardless of the correctness of the fraud accusations. Parents and caregivers were requested to provide additional evidence to prove their entitlement to benefits. However, when they tried to find out what information was considered incorrect or false, or what evidence was deemed missing, they were often met with silence; the tax authorities consistently refused to clarify their decisions. It was also impossible for parents and caregivers, journalists, politicians, oversight bodies and civil society to obtain meaningful information about the existence and workings of the risk classification model.

As a result, a national scandal ensued in which tens of thousands of parents and caregivers were falsely accused of childcare benefit fraud by the Dutch tax authorities. The Dutch childcare benefits scandal was brought to public attention in 2018 and is an ongoing political scandal in the Netherlands today. It led to the fall of the Dutch Cabinet in 2021. The scandal involved improper government action, including but not limited to harsh rules and policies, rigid interpretations of laws, unjustified accusations of fraud, ruthless benefits recovery policies, obstruction of legal and investigative procedures, inadequate and incorrect information, opacity of the childcare benefits fraud system, and the failure of the Dutch authorities to take action in response to people who raised the alarm.

Parents and caregivers who acted in good faith and accidentally made minor administrative errors on applications or renewals were wrongfully accused of fraud. A missing signature on the childcare services contract or a late or incomplete payment of the mandatory personal contribution had severe consequences. Parents and caregivers had to repay large sums of money (payable immediately and in a lump sum) and were labelled as fraudsters. This led to devastating financial problems for the families, ranging from debt

and unemployment to forced evictions because people were unable to pay their rent or make payments on their mortgages. Others were left with mental health issues and stress on their personal relationships, leading to divorces and broken homes.

The scandal also included racial profiling by the risk classification model, which is the focus of this report. The tax authorities used information on whether an applicant had Dutch nationality as a risk factor in the algorithmic system. “Dutch citizenship: yes/no” was used as a parameter in the risk classification model for assessing the risk of inaccurate applications. Consequently, people of non-Dutch nationalities received higher risk scores. The use of the risk classification model amounted to racial profiling. The use of nationality in the risk classification model reveals the assumptions held by the designer, developer and/or user of the system that people of certain nationalities would be more likely to commit fraud or crime than people of other nationalities. It is also indicative of the tax authorities’ perception that there is a link between race/ethnicity and crime, as well as an acceptance of the practice of generalizing the behaviour of some individuals to all others who are perceived to share the same race or ethnicity.

Nationality was used as a constitutive element to pinpoint certain societal groups, based on the idea that these groups supposedly shared certain common cultural values, traditions and/or backgrounds that would make them more likely to engage in fraudulent or criminal behaviour. Using nationality as a factor in risk-scoring for law enforcement purposes in the search for potential perpetrators, or crime or fraud suspects, as was the case with the risk classification model, amounts to differential treatment based on ethnicity. This treatment of parents and caregivers with non-Dutch nationalities by the Dutch tax authorities resulted in racial profiling. Different layers of inequality were created through discrimination based on people’s race and ethnicity in the use of the risk classification model, combined with the policies and practices that more often affected people with a lower economic status, who frequently belonged to an ethnic minority. The childcare benefits scandal must therefore be understood through the lens of intersectional discrimination.

The risk classification model was a black box system that included a self-learning algorithm. Black box systems are algorithmic systems whose inputs and workings are not visible to the users of the system or to other parties. The civil servant (as the system user) did not have access to any details about what information had been used as the basis for assigning a specific risk score to an applicant. The self-learning algorithm gave the risk classification model the ability to learn from experiences over time, independently and autonomously, and to make changes to how it worked without these changes being explicitly programmed by the programmers from the tax authorities. The fact the tax authorities used a black box system and a self-learning algorithm obstructed accountability and transparency and were incompatible with the principles of good governance, legality, and the rule of law. The use, workings and effect of the risk classification model were hidden from the public for a long time.

Moreover, due to opacity from the tax authorities, as well as the lack of implementation and enforcement and/or the limited scope of transparency obligations stemming from international data protection law, parents and caregivers were deprived of meaningful information about the results of the risk classification model in their individual cases. They were unable to defend themselves against the system. This opacity obstructed the right to an effective remedy and violated the principle of good governance and the rule of law.

The Dutch tax authorities did not assess the human rights risks and impacts prior to using the risk classification model and did not mitigate any risks to prevent discrimination stemming from the use of the risk classification model. Oversight over the tax authorities’ use of the risk classification model failed, due

to the fragmented nature of the oversight, a lack of binding human rights oversight mechanisms, and the underlying opacity of the tax authorities.

While the Dutch government publicly disapproves of racial profiling, it continues to allow the use of ethnicity and other prohibited grounds of discrimination risk factors as a basis for suspicion and for decision-making in law enforcement. The lack of transparency and accountability of the tax authorities drove the victims of the childcare benefits scandal into a vortex of bureaucratic and opaque legal procedures, with no effective remedies available to the parents and caregivers that were discriminated against with the risk classification model. In January 2021, the Dutch Cabinet announced a fixed amount of compensation for victims. The proposed compensation scheme does not take into account individual assessments and lacks any effective redress for the discrimination victims suffered as a result of the use of the algorithmic decision-making system.

Amnesty International calls on the Dutch government to:

- Provide compensation for discrimination on the basis of nationality, ethnicity, social origin and racial profiling to the parents and caregivers that were selected by the risk classification model partly on the basis of their non-Dutch nationality.
- Refrain from using black box systems in high-risk contexts, such as in the context of social security fraud detection.
- Ban the use of self-learning algorithms in the public sector at the point of deployment for a) decision-making that results in legal effect, b) decision-making and actions that affect the rights and freedoms of individuals, and c) decision-making and actions that have a major impact on society.
- Create the maximum possible transparency regarding public sector use of algorithmic decision-making systems by creating a public registry which includes detailed and comprehensive information of public sector use of algorithmic decision-making systems.
- Be fully transparent and provide meaningful information to affected individuals about the underlying logic, importance and expected consequences of decisions, even if they are not fully automated, regardless of the level of human involvement in the decision-making process.
- Implement a mandatory and binding human rights impact assessment in public sector use of algorithmic decision-making systems, including by law enforcement authorities. This impact assessment must be carried out during the design, development, use and evaluation phases of algorithmic decision-making systems.
- Establish a clear, unambiguous and legally binding ban on the use of data regarding nationality and ethnicity, or proxies thereof, in risk-scoring for law enforcement purposes in search of potential crime or fraud suspects.
- Establish comprehensive and independent human rights oversight mechanisms over the use of algorithmic decision-making systems, which includes oversight over the impacts of such systems on civil and political as well as social and economic rights, in order to strengthen accountability mechanisms and increase human rights protection.

# 3. SCOPE AND METHODOLOGY

The situation in the Netherlands known as the “childcare benefits scandal” refers to a collection of harmful policies and practices of the Dutch tax authorities and the subsequent inadequate responses by other branches of the Dutch government to address the harms of these policies and practices (see Chapter 4). The scandal was caused by an interplay of improper government actions, including harsh rules and policies, rigid interpretations of laws, unjustified accusations of fraud, ruthless benefits recovery policies, discrimination, obstruction of administrative and investigative procedures, inadequate and incorrect information, opacity of the childcare benefits fraud system, and the failure of the Dutch authorities to take action in response to people who raised the alarm, including citizens, whistle-blowers, lawyers, politicians, journalists and the National Ombudsman.<sup>1</sup>

Previous investigations and reports by various institutions have revealed some of the key facts surrounding these practices and policies. However, the full scope of the human rights impacts has yet to be uncovered. This Amnesty International report does not purport to provide a comprehensive overview of all these impacts, nor does it argue that all the harms related to the childcare benefits scandal (directly) relate to the discriminatory practices and policies that are described here. This report highlights a small but important and underreported aspect of the childcare benefits scandal: the risk classification model, a self-learning algorithm used in fraud detection.

To analyse how the childcare benefits algorithm worked, Amnesty International used multiple governmental sources and sources from independent oversight bodies: firstly, the report *Tax Authorities: The processing of the nationality of applicants for childcare benefits*<sup>2</sup> by the Dutch Data Protection Authority, which researched some of the data processing that took place in the context of fraud detection involving childcare benefits by the Dutch tax authorities<sup>3</sup>; secondly, the parliamentary and ministerial documents that were released in the aftermath of the scandal<sup>4</sup>; and thirdly, the reports of the parliamentary debates related to the childcare benefits system.

<sup>1</sup> Netherlands House of Representatives, Verslag - Parlementaire Ondervragingscommissie Kinderopvangtoeslag: Ongekend Onrecht, 17 December 2020, p. 1, [tweedekamer.nl/sites/default/files/atoms/files/20201217\\_eindverslag\\_parlementaire\\_ondervragingscommissie\\_kinderopvangtoeslag.pdf](https://tweedekamer.nl/sites/default/files/atoms/files/20201217_eindverslag_parlementaire_ondervragingscommissie_kinderopvangtoeslag.pdf).

<sup>2</sup> The original title of the report in Dutch is *Belastingdienst/Toeslagen: De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag*.

<sup>3</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020.

<sup>4</sup> The sources will be provided in the footnotes of the main text.

The description of the harms caused by the childcare benefits scandal is based on *Unprecedented Injustice*<sup>5</sup>, the report by the parliamentary investigation committee<sup>6</sup> that investigated the scandal, and media reporting.<sup>7</sup> Amnesty International also wrote to the demissionary<sup>8</sup> Prime Minister of the Netherlands, asking him to respond to our findings.<sup>9</sup> The letter included a description of the childcare benefits scandal, how the childcare benefits algorithmic decision-making system works, the legal framework, and Amnesty International's human rights analysis of the system, as well as detailed findings. The Dutch government responded to this letter on 27 August 2021.<sup>10</sup> Where necessary, Amnesty International has reflected the response of the Dutch government throughout this report. Amnesty International sent the demissionary Prime Minister a second letter including conclusions that were altered based on information Amnesty International received in the first response. The Dutch government responded to this second letter on 15 and 27 September 2021.<sup>11</sup> Where necessary, Amnesty International has reflected these responses in this report as well.

---

<sup>5</sup> The original title of the report in Dutch is *Ongekend Onrecht*.

<sup>6</sup> In Dutch: 'Parlementaire ondervragingscommissie Kinderopvangtoeslag'.

<sup>7</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020.

<sup>8</sup> A demissionary cabinet (or demissionary Prime Minister) is a type of caretaker in the Netherlands that continues the government in the period after the fall or end of a cabinet and before the instalment of a new cabinet.

<sup>9</sup> A copy of this letter was sent to the demissionary Minister of Finance and the demissionary State Secretary of Finance responsible for benefits.

<sup>10</sup> A copy of this letter has been archived by Amnesty International the Netherlands.

<sup>11</sup> A copy of this letter has been archived by Amnesty International the Netherlands.

# 4. BACKGROUND

*This chapter describes the social and political background underlying social benefits fraud detection in the Netherlands and the introduction of the risk classification model by the tax authorities. It also details the effects of being selected by the system, of (wrongly) being labelled a fraudster, and of being accused by the tax authorities of having committed fraud intentionally.*

Around the world, systems of social protection and assistance are increasingly driven by digital data and technologies that are used to automate, predict, identify, detect, and punish people that meet predetermined profiles.<sup>12</sup> Many of these digital systems have been designed with a particular emphasis on rooting out fraud. Fraud is often presented as a major problem that has damaging social and economic consequences.<sup>13</sup> The automation processes of governments are directed more towards controlling people, rather than helping them.<sup>14</sup> During the 1980s and 1990s, the Netherlands experienced an increase in attention on abuse and fraud in the social security system and their supposedly damaging consequences.<sup>15</sup> This discourse was also reflected in public and political debates.<sup>16</sup> As a result of this widespread sentiment, social security enforcement policies in the Netherlands have been increasingly tightened over the years, and the administrative sanctions system revised.<sup>17</sup> The childcare benefits scheme was introduced in 2005.<sup>18</sup> This scheme is enforced by the Dutch tax authorities, who implemented these harsh social security law enforcement policies and administrative sanction system.

<sup>12</sup> UN Special Rapport on extreme poverty and human rights, 11 October 2019, UN Doc. A/74/48037.

<sup>13</sup> See, for example, Maurits Martijn, "Hoe fraude het nieuwe terrorisme werd", *De Correspondent*, 30 April 2015, decorrespondent.nl/2536/hoe-fraude-het-nieuwe-terrorisme-werd/736027228552-a06124e4.

Esther Lammers, "Oud-ombudsman Alex Brenninkmeijer ziet in de toeslagenaffaire geen bedrijfsongeluk maar een falend systeem", *Trouw*, 31 December 2020, trouw.nl/politiek/oud-ombudsman-alex-brenninkmeijer-ziet-in-de-toeslagenaffaire-geen-bedrijfsongeluk-maar-een-falend-systeem.

For example, with reference to fraud with basic pension amounting to only 0.0001% of the total amount of basic pension paid to claimants, the National Ombudsman at that time stated in 2013 that there is less fraud with social security benefits in the Netherlands than claimed by the government. See "Ombudsman: beeld omvang fraude is verkeerd", *Het Parool*, 11 December 2013, parool.nl/nieuws/ombudsman-beeld-omvang-fraude-is-verkeerd-b5678e0a.

<sup>14</sup> UN Special Rapport on extreme poverty and human rights, 11 October 2019 (previously cited), para. 32.

<sup>15</sup> Goudswaard and others, "Handhaving in de sociale zekerheid", in Michiels and Muller (editors), *Handhaving: Bestuurlijk handhaven in Nederland*, 2014, pp. 565-581, scholarlypublications.universiteitleiden.nl/access/item%3A2915577/view.

<sup>16</sup> See for example Lukas van der Storm, "Wantrouwen en willekeur: over de vlijmscherpe randen van de fraudewet", *Trouw*, 2 February 2021, trouw.nl/verdieping/wantrouwen-en-willekeur-over-de-vlijmscherpe-randen-van-de-fraudewet-b5730f4b.

Chris van der Heijden, "Wantrouwen als leidraad: Een labyrint van protocollen", *De Groene Amsterdammer*, 8 January 2020, groene.nl/artikel/een-labyrint-van-protocollen.

<sup>17</sup> Goudswaard and others, "Handhaving in de sociale zekerheid", in Michiels and Muller (editors), *Handhaving: Bestuurlijk handhaven in Nederland*, 2014, pp. 565-581, scholarlypublications.universiteitleiden.nl/access/item%3A2915577/view.

<sup>18</sup> Algemene wet inkomensafhankelijke regelingen, 23 June 2005.

Parents and caregivers must cover the costs of childcare for their children. However, the Dutch government reimburses a portion of these costs, based on their income. Every parent or caregiver who legally resides in the Netherlands while working or studying is eligible for childcare benefits when their child goes to a registered day-care centre.<sup>19</sup> The distribution of costs between the government and the parent or caregiver depends on various factors, including total family income.<sup>20</sup> For example, parents or caregivers in a low-income household can be reimbursed for up to 96% of their childcare costs through the benefit scheme.<sup>21</sup> Families who have a high income are reimbursed 33.3% of childcare costs.<sup>22</sup> The benefits are paid in advance; the final amount the parent or caregiver is entitled to is calculated afterwards, resulting in a possible surplus amount that must be repaid.<sup>23</sup>

Fearful of people with malicious intentions gaming the social security system, from 2011 onwards, the Dutch government embraced increasingly harsh and highly automated law enforcement policies and practices aimed at combating childcare benefits fraud, leaving little to no room to adapt measures to the individual situation of parents and caregivers.<sup>24</sup> What followed was a “tough on crime” approach that was very harsh on parents and caregivers – even in cases where they had done nothing wrong, had only made minor errors, or had merely made an administrative omission.

Moreover, in 2013 the tax authorities adopted an algorithmic decision-making system for fraud detection in childcare benefits applications and renewals. Known as the risk classification model, this system included self-learning elements<sup>25</sup> to create risk profiles of childcare benefits applicants who were supposedly more likely to submit inaccurate and false information and potentially commit fraud.<sup>26</sup> Parents and caregivers who were selected by the risk classification model due to a high-risk score were subsequently investigated by civil servants. The immediate effect of selection by the risk classification model was suspension of payments of benefits.<sup>27</sup> Chapter 5 describes in detail how this system works and the role that the nationality of the parent or caregiver played in this context.

The tax authorities had to prove the efficiency of their data-driven fraud detection methods, including the new risk classification model, by seizing enough funds from alleged fraudsters to cover the costs of the new methods.<sup>28</sup> This challenge created a perverse incentive to seize as many funds as possible, regardless

<sup>19</sup> Algemene wet inkomensafhankelijke regelingen, 23 June 2005, Article 9(2) in conjunction with Vreemdelingenwet, Article 11. Wet kinderopvang Article 1.1. See also the information about childcare allowance provided on the website of the Dutch government: [rijksoverheid.nl/onderwerpen/kinderopvangtoeslag](http://rijksoverheid.nl/onderwerpen/kinderopvangtoeslag).

<sup>20</sup> Wet kinderopvang, Chapter 5. Besluit kinderopvangtoeslag, Articles 3 – 8.

<sup>21</sup> Besluit kinderopvangtoeslag, Article 6, Annex I.

<sup>22</sup> Besluit kinderopvangtoeslag, Article 6, Annex I.

<sup>23</sup> Algemene wet inkomensafhankelijke regelingen, Article 16.

<sup>24</sup> Netherlands House of Representatives, Report: Verslag - *Parlementaire Ondervragingscommissie Kinderopvangtoeslag: Ongekend Onrecht*, 17 December 2020, pp. 7 and 14, [tweedekamer.nl/sites/default/files/atoms/files/20201217\\_eindverslag\\_parlementaire\\_ondervragingscommissie\\_kinderopvangtoeslag.pdf](http://tweedekamer.nl/sites/default/files/atoms/files/20201217_eindverslag_parlementaire_ondervragingscommissie_kinderopvangtoeslag.pdf).

This is a report about the investigation led by the parliamentary investigative committee on childcare benefits, which was appointed by the Dutch parliament and tasked with investigating the problems surrounding government fraud policies in regard to childcare benefits.

<sup>25</sup> Chapter 2 defines self-learning algorithms and outlines the self-learning elements in the algorithm used by the Dutch tax authorities.

<sup>26</sup> Dutch Data Protection Authority, *Belastingdienst/Toeslagen: De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag*, 16 July 2020, p. 14, [autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek\\_belastingdienst\\_kinderopvangtoeslag.pdf](http://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek_belastingdienst_kinderopvangtoeslag.pdf).

<sup>27</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 16.

<sup>28</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 16.

of the accuracy of the accusations of fraud.<sup>29</sup> One example of the consequences of a high risk score in the risk classification model was that parents and caregivers were requested by civil servants to provide additional evidence to prove their entitlement to benefits. However, parents and caregivers who tried to find out what information was considered incorrect or false, or what evidence was deemed missing, were often met with silence; the tax authorities consistently refused to clarify their decisions.<sup>30</sup> Parents and caregivers did not stand a chance of proving their innocence for many years.<sup>31</sup> It was also impossible for parents and caregivers, journalists, politicians, oversight bodies and civil society to obtain meaningful information about the existence and workings of the risk classification model.<sup>32</sup>

Civil servants who manually reviewed the high-risk applications selected by the risk classification model could, regardless of the severity of potential errors, flag the parent or caregiver as a fraudster.<sup>33</sup> This label had far-reaching consequences. Parents and caregivers were, for example, forced to pay back all or part of the childcare benefits they had received, easily amounting to tens of thousands of euros.<sup>34</sup> Moreover, a missing signature on the childcare services contract or a late or incomplete payment of the mandatory personal contribution to childcare services was enough for the tax authorities to seize all previously received benefits and discontinue all future benefits.<sup>35</sup>

Parents and caregivers with lower incomes often had to pay back larger sums of money, simply because they had received more benefits due to their financial situation. Worse yet, parents and caregivers who were labelled as fraudsters and had also received more than €10,000 in benefits a year<sup>36</sup> or had to pay back more than €3,000 in total<sup>37</sup> were, by default, also flagged for deliberate intent or gross negligence, without any verification. The assignment of the additional label of “deliberate intent or gross negligence” meant that parents and caregivers were no longer eligible for a payment scheme to repay the received benefits in instalments over time.<sup>38</sup> The tax authorities adopted a blanket approach and used the €3,000 threshold solely to expedite decision-making on thousands of applications from parents and caregivers to access the payment schemes.<sup>39</sup>

<sup>29</sup> Chris van der Heijden, “De Belastingdienst bepaalde vooraf hoeveel fraudebestrijding moest opleveren”, *Trouw*, 27 March 2020, [trouw.nl/nieuws/de-belastingdienst-bepaalde-vooraf-hoeveel-fraudebestrijding-moest-opleveren](https://trouw.nl/nieuws/de-belastingdienst-bepaalde-vooraf-hoeveel-fraudebestrijding-moest-opleveren). The Economist, “A benefits scandal sinks the Dutch government”, 21 January 2021, [economist.com/europe/2021/01/23/a-benefits-scandal-sinks-the-dutch-government](https://www.economist.com/europe/2021/01/23/a-benefits-scandal-sinks-the-dutch-government).

<sup>30</sup> See, for example, Nationale Ombudsman, *Geen powerplay maar fair play: onevenredig harde aanpak van 232 gezinnen met kinderopvangtoeslag*, 9 August 2017, p. 4, [nationaleombudsman.nl/system/files/onderzoek/Rapport%202017-095%20Geen%20powerplay%20maar%20fair%20play\\_0.pdf](https://nationaleombudsman.nl/system/files/onderzoek/Rapport%202017-095%20Geen%20powerplay%20maar%20fair%20play_0.pdf).

Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 60.

<sup>31</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 8.

<sup>32</sup> See Section 7.3 and 7.5 on the opacity of the algorithm and the failing oversight. Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 8.

<sup>33</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 16. Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 7.

<sup>34</sup> See, for example, the Administrative Jurisdiction Division of the Netherlands Council of State (ABRvS), ECLI:NL:RVS:2017:589, 8 March 2017, para. 4.2 [uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RVS:2017:589](https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RVS:2017:589). Jan Kleinnijenhuis, “Wie wist wat in de toeslagenaffaire? De kluwen van hoofdrolspelers ontward”, *Trouw*, 14 November 2020, [trouw.nl/politiek/wie-wist-wat-in-de-toeslagenaffaire-de-kluwen-van-hoofdrolspelers-ontward](https://trouw.nl/politiek/wie-wist-wat-in-de-toeslagenaffaire-de-kluwen-van-hoofdrolspelers-ontward).

<sup>35</sup> Advisory committee on implementation of benefits, *Omzien in verwondering: Interim-advies*, 14 November 2019, p. 28-29, [tweedekamer.nl/kamerstukken/amendementen/detail?id=2019Z22146&did=2019D46007](https://tweedekamer.nl/kamerstukken/amendementen/detail?id=2019Z22146&did=2019D46007). Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 109. see also Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 16.

<sup>36</sup> This amount can easily be explained, since benefits were granted to cover the costs of several children in day-care for 32 hours a week.

<sup>37</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 26.

<sup>38</sup> Advisory committee on implementation of benefits, *Omzien in verwondering 2: Eindadvies Adviescommissie Uitvoering Toeslagen*, 12 March 2020, p. 48, [rijksoverheid.nl/documenten/kamerstukken/2020/03/12/omzien-in-verwondering-eindadvies-adviescommissie-uitvoering-toeslagen](https://rijksoverheid.nl/documenten/kamerstukken/2020/03/12/omzien-in-verwondering-eindadvies-adviescommissie-uitvoering-toeslagen).

<sup>39</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 26.

Families were trapped in frustrating and opaque procedures that dragged on for years, with some being driven into poverty and bankruptcy by the government.<sup>40</sup> These practices led to devastating financial problems for the families, ranging from increased debt and unemployment to forced evictions when people were unable to pay their rent or make payments on their mortgages.<sup>41</sup> Others were left with mental health issues and stress on their personal relationships, leading to divorces and broken homes.<sup>42</sup>

In 2016 and 2017, four different government advisory councils warned that the way in which the benefits system had been designed could get people into serious financial trouble within a short time frame, and that these problems would lead to stress, illness, and other problems.<sup>43</sup> However, it took years before the scale and seriousness of the scandal were acknowledged by those legally and politically responsible. From parliamentary documents<sup>44</sup> and media reporting,<sup>45</sup> it has become clear that tens of thousands<sup>46</sup> of parents, caregivers and children<sup>47</sup> have been affected by the scandal.<sup>48</sup> In particular, people with an immigrant background and/or from low-income households have been disproportionately affected.<sup>49</sup>

Years later, long after the harms had been inflicted, a parliamentary investigation committee was appointed by the Dutch parliament in July 2020. The committee was tasked with investigating government policies surrounding childcare benefit fraud. In December 2020, the parliamentary investigation committee concluded that the injustice done to victims of these policies was unprecedented.<sup>50</sup> A string of

<sup>40</sup> Jesse Frederik, "Tienduizenden gedupeerden, maar geen daders: zo ontstond de tragedie achter de toeslagenaffaire", *De Correspondent*, 13 November 2020, [decorrespondent.nl/10628/tienduizenden-gedupeerden-maar-geen-daders-zo-ontstond-de-tragedie-achter-de-toeslagenaffaire](https://decorrespondent.nl/10628/tienduizenden-gedupeerden-maar-geen-daders-zo-ontstond-de-tragedie-achter-de-toeslagenaffaire).

<sup>41</sup> Orkun Akinci, "Gedupeerden van de toeslagenaffaire werden gesloopt door de fiscus: 'Ik wil de overheid de kans geven het goed te maken'", *Trouw*, 16 December 2020, [trouw.nl/leven/gedupeerden-van-de-toeslagenaffaire-werden-gesloopt-door-de-fiscus-ik-wil-de-overheid-de-kans-geven-het-goed-te-maken~b4dfc6c4](https://trouw.nl/leven/gedupeerden-van-de-toeslagenaffaire-werden-gesloopt-door-de-fiscus-ik-wil-de-overheid-de-kans-geven-het-goed-te-maken~b4dfc6c4).  
Barbara Vollebregt, "Gedupeerde ouders krijgen hun leven nooit terug, vrezen ze. 'Koop ik vandaag eten of betaal ik de huur-bdeb0e6e'", *Trouw*, 24 April 2021, [trouw.nl/economie/gedupeerde-ouders-krijgen-hun-leven-nooit-terug-vrezen-ze-koop-ik-vandaag-eten-of-betaal-ik-de-huur~bdeb0e6e](https://trouw.nl/economie/gedupeerde-ouders-krijgen-hun-leven-nooit-terug-vrezen-ze-koop-ik-vandaag-eten-of-betaal-ik-de-huur~bdeb0e6e).

NPO Radio 1, "Toeslagenaffaire heeft ook enorme impact op kinderen: stress en vertrouwen kwijt", 9 November 2020, [nporadio1.nl/nieuws/binnenland/46b6c840-14c3-4219-86f6-06bb4dbd547a/toeslagenaffaire-heeft-ook-enorme-impact-op-kinderen-stress-en-vertrouwen-kwijt-en](https://nporadio1.nl/nieuws/binnenland/46b6c840-14c3-4219-86f6-06bb4dbd547a/toeslagenaffaire-heeft-ook-enorme-impact-op-kinderen-stress-en-vertrouwen-kwijt-en).

<sup>42</sup> Orkun Akinci, "Sibel werd ziek van de Toeslagenaffaire en kan en wil er niets meer over horen", *Trouw*, 27 November 2020, [trouw.nl/binnenland/sibel-werd-ziek-van-de-toeslagenaffaire-en-kan-en-wil-er-niets-meer-over-horen~b0b63a37](https://trouw.nl/binnenland/sibel-werd-ziek-van-de-toeslagenaffaire-en-kan-en-wil-er-niets-meer-over-horen~b0b63a37).

<sup>43</sup> Netherlands Court of Audit, *Aanpak problematische schulden*, 30 June 2016, [rekenkamer.nl/publicaties/rapporten/2016/06/30/aanpak-problematische-schulden](https://rekenkamer.nl/publicaties/rapporten/2016/06/30/aanpak-problematische-schulden).  
Nationale Ombudsman, *Burgerperspectief op schuldhulpverlening*, 11 May 2016, [nationaleombudsman.nl/nieuws/onderzoeken/2016050-burgerperspectief-op-schuldhulpverlening](https://nationaleombudsman.nl/nieuws/onderzoeken/2016050-burgerperspectief-op-schuldhulpverlening).  
Council of Public Health and Society, *Eenvoud loont. Oplossingen om schulden te voorkomen*, 26 June 2017, [raadrvs.nl/documenten/publicaties/2017/06/26/eenvoud-loont](https://raadrvs.nl/documenten/publicaties/2017/06/26/eenvoud-loont).  
Netherlands Scientific Council for Government Policy, *Eigen schuld? Een gedragswetenschappelijk perspectief op problematische schulden*, 30 June 2016, [wrr.nl/publicaties/verkenningen/2016/06/30/eigen-schuld-een-gedragswetenschappelijk-perspectief-op-problematische-schulden](https://wrr.nl/publicaties/verkenningen/2016/06/30/eigen-schuld-een-gedragswetenschappelijk-perspectief-op-problematische-schulden).

<sup>44</sup> Netherlands House of Representatives, Permanent Committee on Finance and the Economy, "Hersteloperatie kinderopvangtoeslagen", 1 July 2021, [tweedekamer.nl/debat\\_en\\_vergadering/uitgeleest/hersteloperatie-kinderopvangtoeslagen](https://tweedekamer.nl/debat_en_vergadering/uitgeleest/hersteloperatie-kinderopvangtoeslagen).

<sup>45</sup> See, for example, Philip de Witt Wijnen, "Snelle compensatie heeft hersteloperatie Toeslagenaffaire vertraagd", *NRC*, 2 May 2021, [nrc.nl/nieuws/2021/05/02/snelle-compensatie-heeft-hersteloperatie-vertraagd-a4042085](https://nrc.nl/nieuws/2021/05/02/snelle-compensatie-heeft-hersteloperatie-vertraagd-a4042085).

<sup>46</sup> The exact number of parents and caregivers affected is not known. So far, 42,000 parents and caregivers have come forward as victims of the childcare benefits system. 19,000 of them have already received a compensation of €30,000 or more if the financial damage suffered was higher. National government of the Netherlands, "Kabinet geeft richting aan voor herstelregeling", 22 June 2021, [actueel.nieuws/2021/06/22/kabinet-geeft-richting-aan-voor-herstelregeling](https://actueel.nieuws/2021/06/22/kabinet-geeft-richting-aan-voor-herstelregeling).

<sup>47</sup> In June 2021, approximately 70,000 children had already been compensated by the Dutch government, see NOS, "70.000 gedupeerde kinderen toeslagenaffaire krijgen tegemoetkoming", 22 June 2021, <https://nos.nl/artikel/2386210-70-000-gedupeerde-kinderen-toeslagenaffaire-krijgen-tegemoetkoming>.

<sup>48</sup> Jesse Frederik, "Tienduizenden gedupeerden, maar geen daders: zo ontstond de tragedie achter de toeslagenaffaire", *De Correspondent*, 13 November 2020, [decorrespondent.nl/10628/tienduizenden-gedupeerden-maar-geen-daders-zo-ontstond-de-tragedie-achter-de-toeslagenaffaire](https://decorrespondent.nl/10628/tienduizenden-gedupeerden-maar-geen-daders-zo-ontstond-de-tragedie-achter-de-toeslagenaffaire).

<sup>49</sup> Secretariat of the Council of Ministers, Ministry of General Affairs, *Minutes of the Council of Ministers*, 26 April 2021, p. 7, [static.nrc.nl/2021/notulen/notulen.pdf](https://nrc.nl/2021/notulen/notulen.pdf).  
Caribisch Netwerk, "Veel Caribische Nederlanders slachtoffer toeslagenaffaire", 1 February 2021, [caribischnetwerk.ntr.nl/2021/02/01/veel-caribische-nederlanders-slachtoffer-toeslagenaffaire](https://caribischnetwerk.ntr.nl/2021/02/01/veel-caribische-nederlanders-slachtoffer-toeslagenaffaire).  
Jon Henley, "Dutch government resigns over childcare benefits scandal", *The Guardian*, 15 January 2021, [theguardian.com/world/2021/jan/15/dutch-government-resigns-over-child-benefits-scandal](https://theguardian.com/world/2021/jan/15/dutch-government-resigns-over-child-benefits-scandal).

<sup>50</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 7.

parliamentary hearings on the scandal led to the fall of the Dutch government and the resignation of the entire Cabinet of Ministers of the Netherlands in January 2021.<sup>51</sup> In the aftermath of the parliamentary investigation report, the Cabinet admitted that parents and caregivers with non-Dutch nationalities and/or dual citizenship were more likely to be investigated by the tax authorities.<sup>52</sup>

In response to the report of the parliamentary investigative committee, the Dutch State Secretary for Finance announced that all parents and caregivers wrongly accused of fraud would receive financial compensation of at least €30,000.<sup>53</sup> However, this remedial measure has been hampered by a number of problems.<sup>54</sup> It has been frustrated by a lack of cooperation between the authorities involved, delayed by ever-changing political demands related to the implementation of the remedial measures, and seen as nearly unworkable due to the complexity of the procedure.<sup>55</sup> Some parents and caregivers will have to wait years before they are compensated.<sup>56</sup>

<sup>51</sup> Thomas Erdbrink, "Government in Netherlands resigns after benefits scandal", *The New York Times*, 15 January 2021, [nytimes.com/2021/01/15/world/europe/dutch-government-resignation-rutte-netherlands.html](https://www.nytimes.com/2021/01/15/world/europe/dutch-government-resignation-rutte-netherlands.html).

<sup>52</sup> Prime Minister of the Netherlands, Minister of General Affairs, Letter to Parliament in response to the report "Ongekend Onrecht", 15 January 2021, [rijksoverheid.nl/documenten/kamerstukken/2021/01/15/kamerbrief-met-reactie-kabinet-op-rapport-ongekend-onrecht](https://www.rijksoverheid.nl/documenten/kamerstukken/2021/01/15/kamerbrief-met-reactie-kabinet-op-rapport-ongekend-onrecht).

<sup>53</sup> Yvonne Hofs, "Alle gedupeerde ouders in toeslagenaffaire krijgen 30 duizend euro compensatie", *Volkskrant*, 22 December 2020, [volkskrant.nl/nieuws-achtergrond/alle-gedupeerde-ouders-in-toeslagenaffaire-kinderopvang-krijgen-30-duizend-euro-compensatie](https://www.volkskrant.nl/nieuws-achtergrond/alle-gedupeerde-ouders-in-toeslagenaffaire-kinderopvang-krijgen-30-duizend-euro-compensatie).

<sup>54</sup> Mark Lievisse Adriaanse and Philip de Witt Wijnen, "Compensatie gedupeerde ouders dreigt vast te lopen in onuitvoerbaarheid", *NRC*, 6 July 2021, [nrc.nl/nieuws/2021/07/06/compensatie-gedupeerde-ouders-dreigt-vast-te-lopen-in-onuitvoerbaarheid](https://www.nrc.nl/nieuws/2021/07/06/compensatie-gedupeerde-ouders-dreigt-vast-te-lopen-in-onuitvoerbaarheid). State Secretary for Finance, "Aanbieding doorlichting UHT en ongevraagd advies van de BAK", 29 June 2021, [rijksoverheid.nl/documenten/kamerstukken/2021/06/29/aanbieding-doorlichting-uht-en-ongevraagd-advies-van-de-bak](https://www.rijksoverheid.nl/documenten/kamerstukken/2021/06/29/aanbieding-doorlichting-uht-en-ongevraagd-advies-van-de-bak). Andersson Elffers Felix, *Doorlichting Uitvoering Hersteloperatie Toeslagen: Periode februari-juni 2021*, 28 June 2021, [rijksoverheid.nl/documenten/rapporten/2021/06/28/eindrapport-doorlichting-hersteloperatie-toeslagen-uht](https://www.rijksoverheid.nl/documenten/rapporten/2021/06/28/eindrapport-doorlichting-hersteloperatie-toeslagen-uht).

<sup>55</sup> Mark Lievisse Adriaanse and Philip de Witt Wijnen, previously cited, 6 July 2021. State Secretary for Finance, previously cited, 29 June 2021. Andersson Elffers Felix, previously cited, 28 June 2021.

<sup>56</sup> Andersson Elffers Felix, previously cited, 28 June 2021. NOS, "Van Huffelen: afhandeling toeslagenaffaire kan nog 'heel lang duren'", 1 October 2021, <https://nos.nl/nieuwsuur/artikel/2400005-van-huffelen-afhandeling-toeslagenaffaire-kan-nog-heel-lang-duren>.

# 5. DESIGN OF THE ALGORITHMIC DECISION-MAKING SYSTEM

*This chapter describes the risk classification model of the childcare benefits system, how it worked, the way in which data on nationality was used by the risk classification model, and the discriminatory nature of other uses of nationality data for fraud detection in the childcare benefits scandal.*

Over the past few decades, the Dutch tax authorities have increasingly automated the provision of complex public services.<sup>57</sup> In addition to the automated assessment of applications and renewals for granting, paying and reclaiming childcare benefits, the tax authorities monitor all applications to detect inaccuracy and potential fraud.<sup>58</sup> Amongst other tools, a risk classification model is used for this monitoring, which includes a self-learning algorithm to estimate the risk of inaccurate applications and renewals.<sup>59</sup> Self-learning algorithms give algorithmic systems the ability to independently and autonomously learn from experiences over time, and to make changes to how they work without these changes being explicitly programmed by humans. Childcare benefits applications that were given a low-risk score by the system were automatically approved by the tax authorities.<sup>60</sup>

The term “fraud monitoring” is not used by the tax authorities themselves to describe the risk classification model. They refer to monitoring for inaccuracy and for supervision (in Dutch: *toezicht*).<sup>61</sup> However, according to the tax authorities’ own policy documents from 2016, “supervision” was seen as an enforcement instrument at the disposal of the tax authorities to monitor applicants’ compliance with the law and to identify cases of errors or abuse/fraud at an early stage.<sup>62</sup> Therefore, Amnesty International considers the use of the risk classification model by the tax authorities to be a form of fraud monitoring.

The practice of detecting inaccuracies in applications was used as an early warning system for potential fraud. Several dozen risk factors were also used in the design of the risk classification model.<sup>63</sup> If an application met the criteria for a risk factor, it would be awarded one or more points, depending on the

<sup>57</sup> Netherlands House of Representatives, “Plan van aanpak ‘Vereenvoudigingsoperatie Belastingdienst’”, 8 June 2007, [parlementairemonitor.nl/9353000/1/9vvij5epmj1ey0/vi3bb0qxfx6](http://parlementairemonitor.nl/9353000/1/9vvij5epmj1ey0/vi3bb0qxfx6).

<sup>58</sup> Dutch Data Protection Authority, *Belastingdienst/Toeslagen: De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag*, 16 July 2020, pp. 14-15, [autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek\\_belastingdienst\\_kinderopvangtoeslag.pdf](http://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek_belastingdienst_kinderopvangtoeslag.pdf).

<sup>59</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 14.

<sup>60</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 14.

<sup>61</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 14.

<sup>62</sup> Dutch Tax Administration, *Beleidsdoorlichting Uitvoering Toeslagen door Belastingdienst*, October 2016, pp. 16 and 19, [rijksoverheid.nl/documenten/rapporten/2016/12/15/bijlage-beleidsdoorlichting-uitvoering-toeslagen-door-belastingdienst](http://rijksoverheid.nl/documenten/rapporten/2016/12/15/bijlage-beleidsdoorlichting-uitvoering-toeslagen-door-belastingdienst).

<sup>63</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 14.

weight that was given to the risk factor.<sup>64</sup> An applicant was awarded a certain number of points for each risk factor, resulting in a total score for each incoming benefits application: the risk score.<sup>65</sup>

The risk classification model was developed by comparing historic examples of correct and incorrect applications.<sup>66</sup> In practice, this meant that the more an incoming application resembled an application that had previously been classified as inaccurate, the higher the risk score assigned to that incoming application. The system selected applications with the highest risk scores for manual review by a civil servant. The civil servant in turn had the power to label an application as fraudulent.<sup>67</sup> The civil servant, however, was given no information as to why the system had given the application a high-risk score for inaccuracy.<sup>68</sup>

The nationality of the applicant was used to determine whether that person was a Dutch citizen. “Dutch citizenship: yes/no” was used as a parameter in the risk classification model for assessing the risk of inaccurate applications.<sup>69</sup> In response to the findings in this report, the Dutch government has argued that the fact citizenship was considered does not mean that nationality was used as a risk factor.<sup>70</sup> However, this argument has no merit; scoring and selecting people entirely or partly on the basis of whether they do or do not have Dutch citizenship constitutes risk-scoring entirely or partly on the basis of nationality. The label “Dutch citizenship: yes/no” contributes to the risk score and treatment that people receive from the tax authorities. The fact that the risk classification model did not differentiate between the different nationalities that were represented in the group of non-Dutch citizens does not alter the fact that the model made a distinction on the basis of nationality.

In addition to the use of applicants’ nationality in the risk classification model for fraud monitoring purposes, the tax authorities also used nationality in various other ways for detection of alleged *organized* fraud by people whose application had been approved.<sup>71</sup> For example, after suspicions of irregularities and fraud committed by some people with roots or links to a particular country,<sup>72</sup> civil servants searched for other people who had those nationalities by querying their databases, and subsequently retrieved additional information regarding *all* applicants with that nationality.<sup>73</sup> For example, a fraud alert concerning 120 to 150 individuals with Ghanaian nationality resulted in an investigation of *all* 6047 applicants with Ghanaian nationality.<sup>74</sup>

Referencing their previous experiences, the tax authorities used various data – including nationality, family ties, and/or living conditions – to identify larger groups that were classified as one homogeneous population, and thus the entire population was under suspicion of organized fraud.<sup>75</sup> Furthermore,

<sup>64</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 15.

<sup>65</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 15.

<sup>66</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 14.

<sup>67</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 15.

<sup>68</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 15.

<sup>69</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, pp. 16-17.

<sup>70</sup> Response of the Dutch government to ‘Line of response childcare benefits scandal report Amnesty International’, 27 August 2021.

<sup>71</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 19.

<sup>72</sup> Based on the available information, the Dutch Data Protection Authority could not rule out that applicants’ dual nationalities were also retrieved. See Dutch Data Protection Authority, 16 July 2020, p. 21.

<sup>73</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, pp. 20 and 26.

<sup>74</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 26.

<sup>75</sup> In its report, the Dutch Data Protection Authority states the following: “The Dutch Data Protection Authority notes on the basis of the available information that the Benefits division of the tax authority and the Multidisciplinary Team for the Detection of Tax Fraud, acting in response to an indication of fraud, requested the nationality of applicants by submitting a query, received the nationalities in an Excel spreadsheet, and aggregated the data in a table in a quick-scan to determine the homogeneity of the population under investigation. According to the Benefits division, the involvement of individuals from the same living environment could be an indication of organized abuse [of the system]. Such homogeneity could for example be evidenced by familial ties or by having the same nationality.” [translation: Amnesty International, see Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 29].

civil servants spoke in denigrating terms about families with Caribbean roots, referring to them as an “Antillean nest”.<sup>76</sup> Civil servants flagged these applicants through manual selection. This suggests that the tax authorities assumed a link between potential fraud and certain groups of people identified via their nationalities and ethnic backgrounds.

People from a non-Dutch background were seen as more prone to commit fraud, indicative of the tax authorities’ perception that a link exists between race, ethnicity and crime, and an acceptance of the practice of applying generalizations to the behaviour of individuals who share the same race or ethnicity. This practice not only reflects the tax authorities’ attitude towards certain nationalities and ethnic minorities who were negatively stereotyped as deviant or fraudulent, but also drives further stigmatization of these groups,<sup>77</sup> which is incompatible with the prohibition of racial discrimination, as will be outlined in Section 6.2.

---

<sup>76</sup> Jan Kleinnijenhuis, “OM: Geen strafrechtelijk onderzoek naar Belastingdienst in verband met toeslagenaffaire”, *Trouw*, 7 January 2021, [trouw.nl/politiek/om-geen-strafrechtelijk-onderzoek-naar-belastingdienst-in-verband-met-toeslagenaffaire](https://trouw.nl/politiek/om-geen-strafrechtelijk-onderzoek-naar-belastingdienst-in-verband-met-toeslagenaffaire).

<sup>77</sup> UN Special Rapporteur on minority issues, 5 January 2015, UN Doc. A/HRC/28/64.

# 6. DISCRIMINATION BY DESIGN

*The childcare benefits scandal exposed how the risk classification model used by the tax authorities reinforced and exacerbated existing inequalities. It resulted in discrimination and racial profiling of parents and caregivers from minority racial and ethnic groups. The following sections will describe the international legal framework and how the practices of the Dutch tax authorities violated international norms regarding non-discrimination and racial profiling.*

## 6.1 LEGAL FRAMEWORK FOR NON-DISCRIMINATION AND RACIAL PROFILING

One of the most frequently reported impacts of algorithms on human rights is the impact on the right to equality and non-discrimination.<sup>78</sup> Algorithms have repeatedly been shown to perpetuate, amplify and entrench historic discrimination or other biases.<sup>79</sup> This is particularly pertinent when used in law enforcement.<sup>80</sup> Biases generally stem from data imbued with historical biases or through the (deliberately or unconsciously) biased choices of the individuals who design, develop, and deploy algorithmic decision-making systems.<sup>81</sup>

The principles of equality and non-discrimination are among the cornerstones of international human rights protection, enshrined in various human rights treaties.<sup>82</sup> In international human rights law, discrimination is defined as:

<sup>78</sup> Catelijne Muller, “The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law: draft d.d. March 20, 2020”, *ALLAI*, 20 March 2020, p. 10, [allai.nl/wp-content/uploads/2020/06/The-Impact-of-AI-on-Human-Rights-Democracy-and-the-Rule-of-Law-draft.pdf](https://allai.nl/wp-content/uploads/2020/06/The-Impact-of-AI-on-Human-Rights-Democracy-and-the-Rule-of-Law-draft.pdf).

Frederik Zuiderveen Borgesius, “Discrimination, Artificial Intelligence, and Algorithmic Decision-Making”, 2018, p. 9, [rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73](https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73).

<sup>79</sup> Catelijne Muller, previously cited, p. 10.

<sup>80</sup> Catelijne Muller, previously cited, p. 5. Amnesty International and Access Now, *The Toronto Declaration*, 16 May 2018, para. 20, [torontodeclaration.org](https://torontodeclaration.org).

<sup>81</sup> UN Special Rapporteur on extreme poverty and human rights, 11 October 2019, *A/74/48037*, p. 20.

Amnesty International and Access Now, *The Toronto Declaration*, 16 May 2018, para. 20, [torontodeclaration.org](https://torontodeclaration.org).

<sup>82</sup> See, for example, Universal Declaration on Human Rights, Article 7. International Covenant on Civil and Political Rights, Article 26. European Convention on Human Rights, Article 14. Protocol 12 to the European Convention for the Protection of Human Rights and Fundamental Freedoms, Article 1. Charter of Fundamental Rights of the European Union, Article 21.

**“...any distinction, exclusion, restriction or preference which is based on any grounds such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status, and which has the purpose or effect of nullifying or impairing the recognition, enjoyment or exercise by all persons, on an equal footing, of all rights and freedoms”.**

UN Human Rights Committee, General Comment 18, 10 November 1989, para. 7.

Some people suffer from additional or unique forms of discrimination because of multiple characteristics that are part of their identity, or perceived in that context, such as race and social origin.<sup>83</sup> This is referred to as intersectional discrimination.<sup>84</sup> International law obliges states to take measures to eliminate all forms of discrimination.<sup>85</sup>

In non-discrimination law, nationality and ethnicity are generally listed as separate protected grounds.<sup>86</sup> In practice, nationality can also be understood as a constitutive element of ethnicity, which can be defined as the idea that societal groups can be delineated on the basis of common nationality, religious faith, language, cultural and traditional origins and backgrounds.<sup>87</sup> Differential treatment of individuals based on nationality in a way that undermines their human rights is prohibited. However, a distinction between nationals and non-nationals can be justified in very specific situations.<sup>88</sup> For example, states may consider nationality in providing access to welfare and social security because financial aid for individuals is reserved for those who have sufficiently close ties to that state.<sup>89</sup> In those cases, it can be justified to use an individual's nationality in assessing the degree of connection required during the benefits application process.<sup>90</sup>

<sup>83</sup> Council of Europe, “Intersectionality and Multiple Discrimination”, undated, [coe.int/en/web/gender-matters/intersectionality-and-multiple-discrimination](http://coe.int/en/web/gender-matters/intersectionality-and-multiple-discrimination).

<sup>84</sup> E. Tendayi Achiume, “Putting racial equality onto the global human rights agenda”, 2018, *SUR*, Volume 28, [sur.conectas.org/en/putting-racial-equality-onto-the-global-human-rights-agenda](http://sur.conectas.org/en/putting-racial-equality-onto-the-global-human-rights-agenda).

<sup>85</sup> European Convention for the Protection of Human Rights and Fundamental Freedoms, Article 14. Protocol 12 to the European Convention for the Protection of Human Rights and Fundamental Freedoms, Article 1. EU Charter of Fundamental Rights, Article 21. International Convention on the Elimination of All Forms of Racial Discrimination, Article 2(1). International Covenant on Civil and Political Rights, Article 26.

<sup>86</sup> See, for example, European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European non-discrimination law*, July 2010, [fra.europa.eu/sites/default/files/fra\\_uploads/1510-fra-case-law-handbook\\_en.pdf](http://fra.europa.eu/sites/default/files/fra_uploads/1510-fra-case-law-handbook_en.pdf).

<sup>87</sup> European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European non-discrimination law: 2018 edition*, 2018, pp. 198-199, [fra.europa.eu/sites/default/files/fra\\_uploads/fra-2018-handbook-non-discrimination-law-2018\\_en.pdf](http://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-handbook-non-discrimination-law-2018_en.pdf). The Court of Justice of the European Union (CJEU) offers the following definition: “The concept of ethnicity, which has its origin in the idea of societal groups marked in particular by common nationality, religious faith, language, cultural and traditional origins and backgrounds.” See CJEU, *CHEZ Razpredelenie Bulgaria ADt. Komisia za zashtita ot diskriminatsia*, Application C-83/14, 16 July 2015, para. 46, [curia.europa.eu/juris/liste.jsf?num=C-83/14](http://curia.europa.eu/juris/liste.jsf?num=C-83/14).

The European Court of Human Rights (ECtHR) provides a similar definition: “Ethnicity has its origin in the idea of societal groups marked by common nationality, tribal affiliation, religious faith, shared language, or cultural and traditional origins and backgrounds.” See ECtHR, *Timishev v. Russia*, Application 55762/00 and 55974/00, 13 March 2006, para. 55, [hudoc.echr.coe.int/eng?i=001-71627](http://hudoc.echr.coe.int/eng?i=001-71627).

<sup>88</sup> See for example, Committee on Economic, Social and Cultural Rights, General Comment 20, 2 July 2009, UN Doc. E/C.12/GC/20, para. 13. International Convention on the Elimination of All Forms of Racial Discrimination, Article 1(3).

<sup>89</sup> CJEU, *Vestische Arbeit Jobcenter Kreis Recklinghausen v. Jovanna García-Nieto and Others*, Application C-299/14, 25 February 2016, [curia.europa.eu/juris/liste.jsf?num=C-299/14](http://curia.europa.eu/juris/liste.jsf?num=C-299/14).

CJEU, *Elodie Giersch and Others v. État du Grand-Duché de Luxembourg*, Application C-20/12, 20 June 2013, [curia.europa.eu/juris/liste.jsf?num=C-20/12](http://curia.europa.eu/juris/liste.jsf?num=C-20/12).

<sup>90</sup> CJEU, *Elodie Giersch and Others v. État du Grand-Duché de Luxembourg*, Application C-20/12, 20 June 2013, para. 37, [curia.europa.eu/juris/liste.jsf?num=C-20/12](http://curia.europa.eu/juris/liste.jsf?num=C-20/12).

This justification cannot be relied upon when nationality or proxies thereof are used as risk factors in risk-scoring models to identify possible fraudulent or criminal behaviour of individuals or groups. The use of nationality in risk-scoring reveals the assumptions held by the designer, developer and/or user of the risk-scoring system that people of certain nationalities would be more likely to commit fraud or crime than people of other nationalities. In risk-scoring, nationality is used as a constitutive element to pinpoint certain societal groups, based on the idea that these groups supposedly share certain common cultural values, traditions and/or backgrounds that would make them more likely to engage in fraudulent or criminal behaviour. Using nationality as a factor in risk-scoring for law enforcement purposes in search of potential perpetrators or crime or fraud suspects amounts to differential treatment based on ethnicity. In 2020, Tendayi Achiume, UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, held that as states increasingly use emerging digital technologies to calculate risk, greater scrutiny of their potential to have a disparate impact on racial or ethnic minorities must be a state priority. Achiume stated:

**“Because digitalization of welfare systems occurs in societies in which groups are marginalized, discriminated against and excluded on a racial and ethnic basis, these systems are almost guaranteed to reinforce these inequities, unless States actively take preventive steps. Without urgent intervention, digital welfare states risk entrenching themselves as discriminatory digital welfare states.”**

UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, 18 June 2020, UN Doc. A/HRC/44/57, para. 42.

Differential treatment of individuals by law enforcement authorities in specific contexts, such as fraud detection or crime prevention, which is not motivated by objective criteria or reasonable justification, and is based on grounds of, for example, national or ethnic origin, is considered racial profiling.<sup>91</sup> The use of data relating to ethnicity and other protected characteristics can be justified only in very limited circumstances, such as on the basis of individualized suspicion based on objective and verified evidence in a criminal investigation.<sup>92</sup> In the most recent recommendations of the UN Committee on the Elimination of Racial Discrimination, racial profiling is defined as:

<sup>91</sup> The concept of racial profiling translates into Dutch as *etnisch profileren* (ethnic profiling). According to the Committee on the Elimination of Racial Discrimination, racial profiling (a) is committed by law enforcement authorities; (b) is not motivated by objective criteria or reasonable justification; (c) is based on grounds of race, colour, descent, national or ethnic origin or their intersection with other relevant grounds, such as religion, sex or gender, sexual orientation and gender identity, disability and age, migration status, or work or other status; (d) is used in specific contexts, such as controlling immigration and combating criminal activity, terrorism or other activities that allegedly violate or may result in the violation of the law. See Committee on the Elimination of Racial Discrimination, General Recommendation 36, 17 December 2020, UN Doc. CERD/C/GC/36, para. 13. Amnesty International defines racial profiling as the use, without objective and reasonable justification, of personal characteristics such as colour, religion, nationality and/or ethnic origin in law enforcement decision-making, not only as a decisive factor but also in combination with other factors. See Amnesty International, “Amnesty International’s observations to the United Nations Committee on the Elimination of Racial Discrimination No. 36 on Preventing and Combating Racial Discrimination”, 27 June 2019, [amnesty.org/download/Documents/IO4006242019ENGLISH.pdf](https://www.amnesty.org/download/Documents/IO4006242019ENGLISH.pdf).

<sup>92</sup> European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European non-discrimination law: 2018 edition*, 2018, p. 93, [fra.europa.eu/sites/default/files/fra\\_uploads/fra-2018-handbook-non-discrimination-law-2018\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-handbook-non-discrimination-law-2018_en.pdf).

**“The practice of police and other law enforcement officials relying, to any degree, on race, colour, descent or national or ethnic origin as the basis for subjecting persons to investigatory activities or for determining whether an individual is engaged in criminal activity”.**

Committee on the Elimination of Racial Discrimination, General Recommendation 36, 17 December 2020, UN Doc. CERD/C/GC/36, p. 4.

Racial profiling is not always blatant, and can happen covertly, particularly where it is programmed into algorithmic decision-making systems. Law enforcement officials using the algorithmic decision-making system do not have to intentionally differentiate in treatment for racial profiling to occur.<sup>93</sup> Racial profiling can be the result of clearly discriminatory attitudes but could also be the result of unconscious bias. Under all circumstances, racial profiling violates the principle of non-discrimination. It leads to the criminalization of certain categories of persons, reinforcing historic stereotypical associations between, for example, fraud and ethnic origin. Racial profiling has harmful consequences because it contributes to the stigmatization of groups, which has a negative impact on the mindset and well-being of the people affected.<sup>94</sup> This type of differential treatment touches on the core of personal dignity.

International human rights standards also prohibit discrimination on the basis of social origin.<sup>95</sup> Social origin refers to a status imputed to an individual by virtue of an inherited or otherwise acquired social, cultural or economic feature.<sup>96</sup> It may relate to an individual's social situation, for example poverty or homelessness, or to the position that an individual has acquired through birth into a particular social class or community, such as those based on ethnicity, religion, descent-based hierarchies, or ideology.<sup>97</sup> Individuals and groups of individuals must not be subjected to arbitrary treatment on account of belonging to a certain economic or social group or stratum within society.<sup>98</sup>

## **6.2 LEGAL ASSESSMENT OF THE ALGORITHMIC DECISION-MAKING SYSTEM**

By using the nationalities of applicants in the manner described in Chapter 5, the Dutch tax authorities violated the affected applicants' rights to equality and non-discrimination. The tax authorities made a distinction between Dutch and non-Dutch applicants by providing different treatment to people with equal eligibility to benefits. Not having Dutch nationality contributed to receiving a higher risk score from the algorithmic decision-making system that was used in fraud detection, resulting in non-Dutch

<sup>93</sup> Amnesty International, “Observations to the United Nations Committee on the Elimination of Racial Discrimination No. 36 on Preventing and Combating Racial Discrimination”, June 2019, [amnesty.org/download/Documents/IO4006242019ENGLISH.pdf](https://www.amnesty.org/download/Documents/IO4006242019ENGLISH.pdf).

<sup>94</sup> See, for example, Ben Bradford and others, “Contact and confidence: revisiting the impact of public encounters with the police”, 18 March 2009, *Policing and Society*, Volume 19, Issue 1, [doi.org/10.1080/10439460802457594](https://doi.org/10.1080/10439460802457594).

<sup>95</sup> International Covenant on Economic, Social and Cultural Rights, Article 2(2).

<sup>96</sup> Committee on Economic, Social and Cultural Rights, General Comment 20, 2 July 2009, UN Doc. E/C.12/GC/20, paras. 24, 25, 26 and 35.

<sup>97</sup> European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European non-discrimination law: 2018 edition*, 2018, p. 218 [fra.europa.eu/sites/default/files/fra\\_uploads/fra-2018-handbook-non-discrimination-law-2018\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-handbook-non-discrimination-law-2018_en.pdf).

<sup>98</sup> European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European non-discrimination law: 2018 edition*, 2018, p. 218, [fra.europa.eu/sites/default/files/fra\\_uploads/fra-2018-handbook-non-discrimination-law-2018\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-handbook-non-discrimination-law-2018_en.pdf).

people being more likely to be automatically selected by the system for additional manual review by a civil servant from the tax authorities. The tax authorities had no objective and reasonable justification for different treatment of non-Dutch applicants and Dutch applicants in fraud detection.

The Dutch government has stated that the Dutch Data Protection Authority did not establish that civil servants terminated some or all benefits based on nationality *after* people were selected by the risk classification model.<sup>99</sup> However, this is a moot point: Amnesty International does not claim that individual civil servants chose to further investigate parents and caregivers on the basis of their nationality, nor that civil servants terminated benefits with discriminatory intent. Instead, Amnesty points out that, prior to investigation by a civil servant, parents and caregivers were selected by the risk classification model partly based on their nationality. As a consequence, people of non-Dutch nationality were more likely to be subjected to additional inspection. In other words: the use of nationality in the risk classification model resulted in differential treatment of individuals based on nationality, without any objective and reasonable justification. This in and of itself constitutes discrimination.

As outlined in Chapter 5, the tax authorities used nationalities for purposes of detecting organized fraud, using nationality to pinpoint certain societal homogenous groups with common cultural values, traditions, and backgrounds, and assuming that these groups are more likely to engage in fraudulent or criminal behaviour than people of other nationalities. The choice to use the criterion of nationality to risk-score inaccuracies in applications for fraud detection reveals the perception within the tax authorities that non-Dutch people in general, and ethnic minorities in particular, are more prone to fraudulent or criminal behaviour. This qualifies as discrimination on the basis of ethnic origin.

The inclusion of nationality as a factor in the risk classification model, in combination with the mapping of certain groups of people who the tax authorities believed would be more likely to engage in fraudulent or criminal behaviour, shows that the tax authorities were motivated by racial prejudice in relation to fraud detection. This risk-scoring led to a disproportionate focus on particular groups of people based on their ethnicity, and qualifies as racial profiling under the international human rights framework. Even in situations where discriminatory policies and practices within the tax authorities were not intentional, the outcome and effect of these policies and practices were discriminatory because they led to differential treatment without an objective and reasonable justification.

When confronted with this analysis, the Dutch government responded that data points on race and ethnicity were not processed by the risk classification model.<sup>100</sup> However, Amnesty International has not claimed that data points on race and ethnicity were processed in the risk classification model. Amnesty International argues that, by processing nationality data for the determination of Dutch citizenship and risk-scoring someone on the basis of whether or not they had Dutch citizenship, the tax authorities discriminated on the basis of ethnicity and conducted racial profiling. The latter finding was not disputed by the Dutch government.

Furthermore, once a parent or caregiver was selected for additional checks, there was a higher risk of falling victim to other improper government actions associated with the childcare benefits scandal, such as harsh rules and policies, rigid interpretations of laws, unjustified accusations of fraud, and ruthless benefits recovery policies.

---

<sup>99</sup> Response of the Dutch government to 'Line of response childcare benefits scandal report Amnesty International', 27 August 2021.

<sup>100</sup> Response of the Dutch government to "Line of response childcare benefits scandal report Amnesty International", 27 August 2021.

These actions had a disproportionate effect on people in low-income brackets in several ways. Firstly, people with low incomes are more dependent on income support and benefits to make ends meet. Benefit cuts and suspensions could lead to immediate financial trouble for parents and caregivers from this group. Secondly, the lower an individual's income, the higher the total benefits that person receives.<sup>101</sup> Low-income earners easily receive thousands of euros in benefits per year.<sup>102</sup> The label of "deliberate intent or gross negligence" was more often (automatically) assigned to people from low-income households, simply because they received more than €10,000 in benefits or had to pay back more than €3,000 to the tax authorities. (See also page 12.) Thirdly, people with low incomes usually have limited flexibility to budget for temporary or permanent suspension of benefits payments, let alone repay amounts already received or pay fines. Repaying benefits and paying fines is often more difficult for this group. The €10,000 and €3,000 thresholds defined in the childcare benefits policy had a disproportionately adverse effect on low-income households that would have benefitted most from the payment scheme, because they were often unable to repay large sums of money at once.<sup>103</sup> The punitive approach adopted towards fraud particularly affected people in lower income brackets, a segment of Dutch society which has high representation of ethnic minorities.<sup>104</sup>

As such, different layers of inequality were created through discrimination based on race and ethnicity in the use of the risk classification model, combined with the policies and practices that more often affected people with a lower economic status, who often belonged to an ethnic minority. The childcare benefits scandal must therefore be understood through the lens of intersectional discrimination.

---

<sup>101</sup> Statistics Netherlands (CBS), 'Recordaantal kinderen met kinderopvangtoeslag', 3 July 2018, [cbs.nl/nl-nl/nieuws/2018/27/recordaantal-kinderen-met-kinderopvangtoeslag](https://www.cbs.nl/nl-nl/nieuws/2018/27/recordaantal-kinderen-met-kinderopvangtoeslag).

<sup>102</sup> The test calculation offered by the tax authorities on their website shows that, for example, a single parent or caregiver receiving minimum wage with a full-time job and two children could receive more than €3,000 in childcare benefits per month. Benefits Division of the Dutch tax authorities, "Proefberekening toeslagen", [belastingdienst.nl/wps/wcm/connect/nl/toeslagen/content/hulpmiddel-proefberekening-toeslagen](https://belastingdienst.nl/wps/wcm/connect/nl/toeslagen/content/hulpmiddel-proefberekening-toeslagen).

<sup>103</sup> Advisory committee on implementation of benefits, *Omzien in verwondering: Interim-advies*, 14 November 2019, p. 48, [tweedekamer.nl/kamerstukken/amendementen/detail?id=2019Z22146&did=2019D46007](https://www.tweedekamer.nl/kamerstukken/amendementen/detail?id=2019Z22146&did=2019D46007).

<sup>104</sup> Bart van Hulst and others, "Armen met en zonder migratieachtergrond", 3 September 2019, Netherlands Institute for Social Research (SCP), [digitaal.scp.nl/armoedein kaart2019/armen-met-en-zonder-migratieachtergrond](https://digitaal.scp.nl/armoedein kaart2019/armen-met-en-zonder-migratieachtergrond).

# 7. FAILURE TO PROTECT HUMAN RIGHTS

*States have the obligation to respect, protect and uphold human rights, including when using algorithmic decision-making systems. By adopting tailored measures and safeguards, such as enhanced transparency, the risk to human rights violations when using such systems can be mitigated. This chapter illustrates the minimum measures and safeguards that should be put in place, and how the lack of these measures contributed to the harms inflicted by the childcare benefits scandal in the Netherlands.*

## 7.1 SOUNDING LIKE A BROKEN RECORD: PRIOR WARNINGS FOR HUMAN RIGHTS RISKS

*The Dutch government received various warnings about the human rights impacts of its use of algorithmic decision-making systems. Despite that, no binding measures or human rights safeguards have been taken to reduce the risk to human rights, despite the rapid roll-out of these systems continuing unabated in Dutch government processes.*

The childcare benefits scandal did not emerge in a vacuum. In 2019, the UN Special Rapporteur on extreme poverty and human rights warned the Netherlands about the human rights risks associated with an algorithmic decision-making system designed to analyse a wide range of personal and sensitive data to predict how likely people are to commit tax fraud or benefits fraud.<sup>105</sup> The system had been exclusively deployed in designated “problem neighbourhoods” across the Netherlands – a potential proxy that led to discrimination risks for people living in poverty and people from immigrant backgrounds.<sup>106</sup> The system was later prohibited by a Dutch court, which considered it unlawful and in violation of human rights.<sup>107</sup> Furthermore, an Amnesty International investigation revealed in 2020 that Roermond, a city in the south of the Netherlands, was being used as a “living lab” by the Dutch National Police to develop risk-scoring models.<sup>108</sup> The living lab was based on mass surveillance and discrimination against people of Eastern European nationalities and/or Roma ethnicity.<sup>109</sup> Some of the concerns and recommendations included in this report are similar to those covered in the 2020 report on the living lab in Roermond.<sup>110</sup>

<sup>105</sup> UN Special Rapporteur on extreme poverty and human rights, 11 October 2019, UN Doc. A/74/493. UN Special Rapporteur on Extreme poverty and human rights, “Brief by the UN Special Rapporteur on extreme poverty and human rights as *Amicus Curiae* in the case of NJCM c.s./De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number C/09/550983/HA ZA 18/388), 26 September 2019, [ohchr.org/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf](https://ohchr.org/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf).

<sup>106</sup> District Court of The Hague, ECLI:NL:RBHDHA:2020:865, 5 February 2020, [uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBHDHA:2020:865](https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBHDHA:2020:865).

<sup>107</sup> District Court of The Hague (previously cited), 5 February 2020.

<sup>108</sup> Netherlands Police, “Zakkenrollers herkennen dankzij data-koppeling”, 11 July 2018, [politie.nl/nieuws/2018/juli/11/00-zakkenrollers-herkennen-dankzij-data-koppeling](https://politie.nl/nieuws/2018/juli/11/00-zakkenrollers-herkennen-dankzij-data-koppeling).

<sup>109</sup> Amnesty International, *We Sense Trouble: Automated Discrimination and Mass Surveillance in Predictive Policing in the Netherlands* (Index: EUR 35/2971/2020), 28 September 2020, [amnesty.org/en/documents/eur35/2971/2020/en](https://amnesty.org/en/documents/eur35/2971/2020/en).

<sup>110</sup> Amnesty International (previously cited), 28 September 2020.

It is an explicit ambition of the Dutch state to be a leader in the field of algorithmic governance, in which ethical values and human rights are safeguarded.<sup>111</sup> However, despite the previous warnings described above, no steps have been taken by the Dutch government to fulfil this ambition, other than to echo that there are serious human rights risks associated with the use of algorithmic decision-making systems.<sup>112</sup>

In response to the findings in this report, the Dutch government explained that there are ongoing improvements with regard to policies for the use of algorithmic decision-making systems.<sup>113</sup> Amnesty International views this as a step in the right direction. However, these efforts are not sufficient to address the human rights concerns that are identified in this report and do not meet Amnesty International's recommendations to protect human rights in the use of algorithmic decision-making systems. For example, no binding obligations have been laid down that require designers, developers and users of algorithmic decision-making systems to identify, prevent, and mitigate potential and actual adverse human rights impacts. Moreover, human rights oversight and accountability and transparency mechanisms for the use of algorithmic decision-making systems in the Netherlands continues to be weak.<sup>114</sup> Even so, the Dutch government continues to roll out new algorithmic decision-making systems in the public sector and laws that facilitate high-speed mass data exchange between governmental bodies for risk-scoring of people.<sup>115</sup>

To prevent similar human rights violations from happening in the future, binding and enforceable measures and safeguards must be implemented. States are the primary duty bearers responsible for protecting human rights.<sup>116</sup> International human rights law applies to the actions of governments, regardless of whether these actions are mediated by algorithmic decision-making systems.<sup>117</sup> Currently, there are no explicit binding international human rights standards for the protection of human rights in algorithmic decision-making systems. The childcare benefits scandal demonstrates the urgent need for states, as well as international, regional, and supranational bodies, to adopt binding rules and standards in order to articulate the obligations and responsibilities to respect, protect and uphold human rights in the design and use of algorithmic decision-making systems.

**Governments must put in place a framework that a) prevents human rights violations in relation to the design and use of algorithmic decision-making systems, b) establish monitoring and oversight mechanisms as safeguards, c) hold those responsible for violations to account, and d) provide effective remedy to individuals and groups whose rights have been violated.**

<sup>111</sup> National Government of the Netherlands, *Nederlandse Digitaliseringsstrategie 2021*, 26 April 2021, [rijksoverheid.nl/documenten/kamerstukken/2021/04/26/nederlandse-digitaliseringsstrategie-2021](https://rijksoverheid.nl/documenten/kamerstukken/2021/04/26/nederlandse-digitaliseringsstrategie-2021).

<sup>112</sup> The Netherlands Court of Audit concluded in a recently published report that algorithms are widely used within the Dutch government, while measures to reduce bias are often lacking in actual practice. General frameworks are not specified for algorithms and are also not applied in a coherent manner. The Court of Audit is an independent body that audits the spending of the national government on its efficiency and legitimacy. The Court of Audit is appointed by the Cabinet on the advice of the House of Representatives. See Netherlands Court of Audit, *Aandacht voor algoritmes*, 26 January 2021, [rekenkamer.nl/publicaties/rapporten/2021/01/26/aandacht-voor-algoritmes](https://rekenkamer.nl/publicaties/rapporten/2021/01/26/aandacht-voor-algoritmes).

<sup>113</sup> Response of the Dutch government to "Line of response childcare benefits scandal report Amnesty International", 27 August 2021. In its response, the Dutch government refers to an assessment framework for algorithms, a human rights impact assessment and a guide for non-discriminatory AI. These instruments are non-binding and dependent on voluntary use. The Dutch government further refers to Cabinet letters which do not contain plans to implement binding human rights safeguards for the use of algorithms, big data and/or risk-profiling.

<sup>114</sup> In addition, the Netherlands Court of Audit concluded that current monitoring of the use of algorithms is inadequate and that citizens are not given insight into their use. See Algemene Netherlands Court of Audit, *Aandacht voor algoritmes*, 26 January 2021, [rekenkamer.nl/publicaties/rapporten/2021/01/26/aandacht-voor-algoritmes](https://rekenkamer.nl/publicaties/rapporten/2021/01/26/aandacht-voor-algoritmes).

<sup>115</sup> Het Parool, "Zeker 25 gemeenten gebruiken algoritmen voor risicoprofilering inwoners", 2 February 2021, [parool.nl/nederland/zeker-25-gemeenten-gebruiken-algoritmen-voor-risicoprofilering-inwoners](https://parool.nl/nederland/zeker-25-gemeenten-gebruiken-algoritmen-voor-risicoprofilering-inwoners).

<sup>116</sup> UN Human Rights Committee, General Comment No. 31, 26 May 2004, UN Doc. CCPR/C/21/Rev.1/Add. 13, paras. 3-8. UN Committee on Economic Social and Cultural Rights, General Comment 3, 14 December 1990, UN Doc. E/1991/23, paras. 2-8.

<sup>117</sup> Committee on the Elimination of Racial Discrimination, General Recommendation No. 36 on preventing and combating racial profiling by law enforcement officials, 17 December 2020, UN Doc. CERD/C/GC/36, para. 58.

## 7.2 OUT OF CONTROL: BLACK BOX SYSTEMS AND SELF-LEARNING ALGORITHMS IN THE PUBLIC SECTOR

*The risk classification model was a black box system that also included a self-learning algorithm. This algorithmic decision-making system obstructed accountability and transparency and was incompatible with the principles of good governance, legality, and the rule of law.*

Black box systems are algorithmic systems whose inputs and workings are not visible to the users of the system or to other parties. These types of systems are opaque and therefore cannot be subjected to meaningful standards of oversight, accountability, and transparency.<sup>118</sup> The risk classification model selected applications with the highest score for inaccuracy and referred them for manual checks and reviews by a civil servant. The civil servant (as the system user) did not have access to any details about what information had been used as the basis for assigning a specific risk score to an applicant.<sup>119</sup> The fact the tax authorities used a black box system obstructed meaningful accountability and transparency of the tax authorities' fraud detection practices and policies.

The risk classification model also included a self-learning algorithm, as explained on page 15. The use of self-learning algorithms in the public sector presents a variety of challenges that pose significant risks to human rights and undermine good governance and the principles of the rule of law. Self-learning algorithms constantly absorb new information and then adapt the way they make decisions. If an algorithm is self-learning, the result can be that at some point no human, including the designers, developers, users, and people working for human rights oversight bodies, understands how a particular decision is made by the system.<sup>120</sup> This is problematic because government activities must be verifiable and predictable. There is also a significant risk that the deployment of black box systems and self-learning algorithms will amplify intentional or unintentional biases, leading to systems that produce results that are systemically prejudiced due to erroneous assumptions embedded in the self-learning process.<sup>121</sup> These characteristics are incompatible with good governance, the principle of legality, and the rule of law when self-learning algorithms are used in the public sector, including in support of decision-making and actions that result in legal effect, affect the rights and freedoms of individuals, or have a major impact on society.<sup>122</sup>

**Governments must refrain from using black box systems in high-risk contexts, such as in social security fraud detection. Governments must ban the use of self-learning algorithms in the public sector at the point of deployment for a) decision-making that results in legal effect, b) decision-making and actions that affect the rights and freedoms of individuals, and c) decision-making and actions that have a major impact on society.**

<sup>118</sup> Amnesty International and Access Now, *The Toronto Declaration*, 16 May 2018, Article 32 sub c, [torontodeclaration.org](https://torontodeclaration.org).

<sup>119</sup> Dutch Data Protection Authority, *Belastingdienst/Toeslagen: De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag*, 16 July 2020, pp. 14-15, [autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek\\_belastingdienst\\_kinderopvangtoeslag.pdf](https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek_belastingdienst_kinderopvangtoeslag.pdf).

<sup>120</sup> Council for the Judiciary, "Algoritmes in de rechtspraak: Wat artificiële intelligentie kan betekenen voor de rechtspraak", 2019, [rechtspraak.nl/SiteCollectionDocuments/rechtstreeks-2019-02.pdf](https://rechtspraak.nl/SiteCollectionDocuments/rechtstreeks-2019-02.pdf).

<sup>121</sup> Amnesty International and Access Now, *The Toronto Declaration*, 16 May 2018, [torontodeclaration.org](https://torontodeclaration.org).

<sup>122</sup> Amnesty International, "Inbreng Amnesty International voor het Commissiedebat Bescherming persoonsgegevens op 20 mei 2021", 17 May 2021, [amnesty.nl/content/uploads/2021/05/Inbreng-Amnesty-commissiedebat-bescherming-persoonsgegevens.pdf?x92023](https://amnesty.nl/content/uploads/2021/05/Inbreng-Amnesty-commissiedebat-bescherming-persoonsgegevens.pdf?x92023).

## 7.3 KEPT IN THE DARK: A LACK OF TRANSPARENCY

*The use, workings and effect of the childcare benefits risk classification model were hidden from the public for a long time. Moreover, due to opacity from the tax authorities, as well as poor implementation and enforcement, or limited scope of transparency obligations stemming from international data protection law, parents and caregivers were deprived of meaningful information about the results of the risk classification model in their individual cases. They were unable to defend themselves against the system. This opacity obstructed the right to an effective remedy and undermined the principle of good governance and the rule of law.*

Transparency is a key principle of good governance and the rule of law.<sup>123</sup> Transparency in the creation and implementation of public policies empowers the public to access social services and demand protection of their rights.<sup>124</sup> Furthermore, the UN Human Rights Committee has expressly acknowledged the general right of access to information held by public bodies.<sup>125</sup> To give effect to this right, states should both proactively publish government information of public interest and “enact the necessary procedures, whereby one may gain access to information, such as by means of freedom of information legislation”.<sup>126</sup> Transparency must be given in a general sense, as well as to individuals who are affected by government decisions and actions.<sup>127</sup>

Given the risk that the use of algorithmic decision-making systems could result in discriminatory outcomes, the need for transparency regarding the use of algorithms in the public sector is even more urgent.<sup>128</sup> In addition, the very existence of algorithmic decision-making systems in public sector decision-making is often hidden or unknown, making it difficult or impossible to know whether it has an impact on human rights in the first place. This results in an asymmetry of information between those negatively impacted by algorithmic decision-making systems and those developing and using such systems, stressing the need to reinforce mechanisms of transparency.<sup>129</sup> Moreover, when transparency of algorithmic decision-making systems is not ensured,<sup>130</sup> the existence of biases can easily remain undetected or even be obscured.<sup>131</sup> Public registries that provide transparency about the use of algorithmic decision-making systems in the

<sup>123</sup> UN, “What is the Rule of Law”, undated, [un.org/ruleoflaw/what-is-the-rule-of-law](http://un.org/ruleoflaw/what-is-the-rule-of-law).

UN Human Rights Office of the High Commissioner, “About Good Governance: What is good governance?”, undated, [ohchr.org/EN/Issues/Development/GoodGovernance/Pages/AboutGoodGovernance.aspx](http://ohchr.org/EN/Issues/Development/GoodGovernance/Pages/AboutGoodGovernance.aspx).

<sup>124</sup> Office of the United Nations High Commissioner for Human Rights, *Good Governance Practices for the Protection of Human Rights*, 2007, p. 5, [ohchr.org/documents/publications/GoodGovernance.pdf](http://ohchr.org/documents/publications/GoodGovernance.pdf).

The Council of Europe Commissioner for Human Rights proposes a series of practical recommendations to protect human rights in algorithmic systems to national authorities on 10 main areas for action, one of which is transparency. See Council of Europe, *Unboxing Artificial Intelligence: 10 steps to protect Human Rights*, May 2019, p. 9, [rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64](http://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64).

<sup>125</sup> International Covenant on Civil and Political Rights, Article 19. UN Human Rights Committee, General Comment 34, 12 September 2011, UN Doc. CCPR/C/GC/34, [ohchr.org/english/bodies/hrc/docs/gc34.pdf](http://ohchr.org/english/bodies/hrc/docs/gc34.pdf).

<sup>126</sup> UN Human Rights Committee, General Comment 34, 12 September 2011, UN Doc. CCPR/C/GC/34, [ohchr.org/english/bodies/hrc/docs/gc34.pdf](http://ohchr.org/english/bodies/hrc/docs/gc34.pdf).

<sup>127</sup> Committee on the Elimination of Racial Discrimination, General Recommendation No. 36 on preventing and combating racial profiling by law enforcement officials, 17 December 2020, UN Doc. CERD/C/GC/36, para. 61.

<sup>128</sup> Amnesty International and Access Now, The Toronto Declaration, 16 May 2018, Article 32, [torontodeclaration.org](http://torontodeclaration.org).

<sup>129</sup> Council of Europe, *Ad Hoc Committee on Artificial Intelligence (CAHAI): Feasibility Study*, 17 December 2020, CAHAI(2020)23, para. 85, [rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da](http://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da).

<sup>130</sup> Committee on the Elimination of Racial Discrimination, General Recommendation No. 36 on preventing and combating racial profiling by law enforcement officials, 17 December 2020, UN Doc. CERD/C/GC/36, para. 64.

<sup>131</sup> Council of Europe, *Ad Hoc Committee on Artificial Intelligence (CAHAI): Feasibility Study*, 17 December 2020, CAHAI(2020)23, para. 30, [rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da](http://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da).

public sector are not yet widely adopted in the Netherlands.<sup>132</sup>

For a long time, the Dutch tax authorities were secretive about the use of the childcare benefits risk classification model and its workings. The tax authorities did not make any general information public about their use of algorithmic decision-making systems, nor the purpose and workings of such systems.<sup>133</sup> As will be explained in Section 7.5., the work of various oversight bodies has been obstructed by the tax authorities' continuing lack of transparency and unwillingness to cooperate.<sup>134</sup> Journalists, politicians, academics, and civil society organizations did not, and still do not, have access to relevant information that is necessary to properly monitor the use of algorithmic decision-making systems and how discriminatory algorithms were affecting families for years on end.<sup>135</sup>

Although the childcare benefit fraud risk classification model was introduced in 2013,<sup>136</sup> it was only in July 2020 that it was officially established by the Dutch Data Protection Authority that applicants' nationality played a role in the risk classification model. If the tax authorities had been more transparent, the harms caused by the discriminatory algorithms they used could have been brought to light at an earlier stage. Even after the childcare benefits scandal was revealed, the Dutch government did not enact a general obligation for transparency around the use of algorithmic decision-making systems. (See the box on page 30 on the recommended specifications of such mechanisms.)

Access to information on how government operations impact human rights is also vital to enable people to claim their right to an effective remedy when their rights have been violated (see also Section 7.5). However, the opacity of some algorithmic decision-making systems and the inability to scrutinize the way these systems work (see Section 7.2 on the "black box phenomenon") means that individuals may be unaware of how decisions affecting their rights were made. It is only when the transparency of an algorithmic decision-making system is ensured, for instance by documentation or logging of relevant information, that a system can be audited, and that it is possible to fully verify the extent to which it may, for instance, infringe on the right to non-discrimination.<sup>137</sup> In particular, where the decisions made by such systems affect marginalized and/or disadvantaged communities, such as people from immigrant backgrounds, extra steps need to be taken to ensure clear, correct and complete information is provided to them in an accessible and understandable way.<sup>138</sup> The right to transparency of algorithmic decision-making systems that process personal data is codified in the Data Protection Convention 108+ of the Council of

<sup>132</sup> In response to a resolution by a Member of Parliament to establish an algorithm registry, the Cabinet is planning a pilot with a transparency tool. However, registries that provide transparency about algorithm use in the public sector are not yet widely adopted and the Cabinet has not proposed plans for transparency obligations. See Letter from the government, "Voortgang AI en algoritmen", 10 June 2021, [tweedekamer.nl/kamerstukken/brieven\\_regering/detail?id=2021Z10411&did=2021D22770](https://tweedekamer.nl/kamerstukken/brieven_regering/detail?id=2021Z10411&did=2021D22770). The city of Amsterdam voluntarily started a pilot with an algorithm registry. The registry does not include all information as recommended on page 30. See <https://algorithmeregister.amsterdam.nl/en/participate-in-a-survey/>. Amnesty International, "Amnesty International Submission to the European Commission's Consultation on Artificial Intelligence", June 2020, Recommendation 5, [ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/public-consultation\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/public-consultation_en).

<sup>133</sup> In January 2021, the Dutch Cabinet acknowledged the criticism on the lack of information and expressed its willingness to be more transparent. The Cabinet announced plans to publish documents that underlie ministerial decisions and internal deliberations. However, transparency on the algorithmic system was not discussed in the plans. See Prime Minister of the Netherlands, Minister of General Affairs, Letter to Parliament in response to the report "Ongekend Onrecht", 15 January 2021, [rijksoverheid.nl/documenten/kamerstukken/2021/01/15/kamerbrief-met-reactie-kabinet-op-rapport-ongekend-onrecht](https://rijksoverheid.nl/documenten/kamerstukken/2021/01/15/kamerbrief-met-reactie-kabinet-op-rapport-ongekend-onrecht).

<sup>134</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 8.

<sup>135</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 8.

<sup>136</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 14.

<sup>137</sup> Amnesty International and Access Now, The Toronto Declaration, 16 May 2018, Article 30 in conjunction with 32, [torontodeclaration.org](https://www.torontodeclaration.org).

<sup>138</sup> UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 30 April 2009, UN Doc. A/ HRC/11/4, paras. 14 and 53.

Europe, of which the Netherlands is a signatory.<sup>139</sup> This right includes access to information about the reasoning underlying the processing of personal data, including the consequences of such reasoning.<sup>140</sup> However, nearly three years after signing Convention 108+, the Netherlands has yet to ratify it.<sup>141</sup>

The parents and caregivers affected by the childcare benefits scandal were kept unaware of the fact that an algorithm had selected their application because it received a high-risk score for inaccuracy. Similarly, parents and caregivers also did not receive general information about the risk factors and their individual scores. Families that belonged to a marginalized group because of their immigration background had a higher risk of being selected (see Chapter 4). The tax authorities undertook no extra steps to ensure clear, correct, and complete information about the algorithmic decision-making system and its decision-making processes. Parents and caregivers had no access to information about the reasoning underlying the processing of their data, including the consequences of such reasoning. Because of this lack of transparency, parents and caregivers were unable to defend themselves or prove their innocence. They were also unable to bring legal claims/challenges on the basis of discrimination in the risk classification model, which obstructed access to justice for victims of the system (see Section 7.7).

Since the tax authorities processed personal data, the obligations laid down in the European Union's General Data Protection Regulation are applicable. Under this regulation, individuals have the right to meaningful information about the logic of algorithms.<sup>142</sup> However, this right only applies in very limited circumstances to decisions made by algorithmic decision-making systems that are fully automated. The right to meaningful information does not extend to algorithmic decision-making systems that assist or support the decision-making of civil servants; a limited or trivial degree of human involvement in decision-making means that the right to meaningful information does not apply.

As explained in Chapter 5, applications that were classified as low risk by the risk classification model were automatically approved by the tax authorities. For these parents and caregivers, the risk classification model functioned as a fully automated decision-making system. This group therefore has the right to meaningful information about the logic of the algorithm behind the risk classification model. Additional scrutiny is needed by means of human intervention in automated or semi-automated decision-making that rejects applications of parents and caregivers, and such scrutiny must complement meaningful information. However, the right to meaningful information is not available to the parents and caregivers

---

<sup>139</sup> Council of Europe, Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, 28 January 1981. The Convention was modernized in 2018. See Council of Europe, Convention 108+, Convention for the protection of individuals with regard to the processing of personal data, 18 May 2018.

<sup>140</sup> Article 9(1)(c) and recital 77 of the Explanatory Report of Convention 108+. It follows from this recital that, for example, in the case of credit scoring, individuals should be entitled to know the logic underpinning the processing of their data and resulting in a "yes" or "no" decision, and not simply information regarding the decision itself. Having an understanding of these elements contributes to the effective exercise of other essential safeguards, such as the right to object and the right to complain to a competent authority.

<sup>141</sup> Council of Europe Portal, "Chart of signatures and ratifications of Treaty 223", 16 September 2021, [coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatyid=223](https://coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatyid=223).

<sup>142</sup> General Data Protection Regulation, Article 22 in conjunction with Articles 12-17.

that were discriminated against. This group was selected for additional scrutiny by a civil servant, meaning that there was human intervention in the decision-making process. Under the General Data Protection Regulation, human intervention substitutes the right to meaningful information. Meaningful information would reveal the logic behind the algorithmic decision-making system and interaction between the system and the civil servant, which helps in detecting incorrect inputs and outputs, discrimination, automation bias, and the use of black box systems.<sup>143</sup> In the childcare benefits scandal, the group to whom transparency matters the most was essentially excluded from access to meaningful information about the logic of the algorithm.

**Governments should create the maximum possible transparency regarding public sector use of algorithmic decision-making systems by creating a public registry which includes detailed and comprehensive information of public sector use of algorithmic decision-making systems. Governments should be fully transparent and provide meaningful information to affected individuals about the underlying logic, importance and expected consequences of decisions, even if they are not *fully* automated, regardless of the level of human involvement in the decision-making process.**

**AN ALGORITHMIC DECISION-MAKING SYSTEM REGISTRY SHOULD CONTAIN AT LEAST THE FOLLOWING INFORMATION:**

- Where and for what purpose(s) the algorithmic decision-making systems are used;
- The assumptions, premise and design principles on which the systems are based;
- The accuracy of the system;
- The identities of the designer, developers and users of the systems, including (commercial) partners and public-private partnerships;
- The data sets that are used to train the model and the purpose of the initial collection of this data;
- How humans utilize predictions that are made by the system;
- What risks and impacts have been identified (see Section 7.4);
- The result of periodic evaluation;
- Any legal and technical safeguards that are in place to mitigate the identified risks;
- What rights the individuals have and information on how individuals can engage these rights;
- What mechanisms exist in order to inform individuals in case something goes wrong;
- How often, by whom and how the system is being evaluated and assessed;
- In what way individuals can give feedback to the designers, developers and users of the algorithmic decision-making system;
- What type of redress individuals can seek.

<sup>143</sup> Moreover, human involvement in algorithmic systems is often subjected to automation bias, severe time pressure or lack of information when reviewing the decision. See for example Kate Goddard and others, "Automation bias: a systemic review of frequency, effect mediators, and mitigators", *Journal of the American Medical Informatics Association*, Volume 19, Issue 1, 16 June 2011, <https://academic.oup.com/jamia/article/19/1/121/732254>.

## 7.4 HUMAN RIGHTS ON THE BACK BURNER: NO ASSESSMENT OF THE HUMAN RIGHTS IMPACTS

*The Dutch tax authorities did not assess the potential human rights risks and impacts prior to using the childcare benefits risk classification model, and therefore did not mitigate the discriminatory impacts stemming from the use of the risk classification model.*

States should continuously assess and monitor the human rights impact of algorithmic decision-making systems throughout their entire lifecycle, and take appropriate mitigation measures if risks or harms to human rights are identified.<sup>144</sup> In a Recommendation on the human rights impacts of algorithmic decision-making systems, the Council of Europe held that states should address all human rights risks by carrying out a human rights impact assessment.<sup>145</sup> The rapidly developing and iterative nature of the technology means that new risks are likely to emerge in a variety of unforeseen contexts.<sup>146</sup> It is therefore vital that all algorithmic decision-making systems are subjected to an ongoing process that ensures the respect, protection and fulfilment of human rights.<sup>147</sup> The onus must be on public sector bodies that are developing and deploying an algorithmic decision-making system to identify the actual and potential impacts on human rights, and demonstrate that they have taken appropriate action to address these risks, on an ongoing, transparent and dynamic basis.<sup>148</sup> A human rights impact assessment for algorithmic decision-making systems can be an important tool to identify potential and actual adverse human rights impacts of algorithmic decision-making systems.<sup>149</sup>

The Dutch tax authorities did not conduct a human rights impact assessment for the childcare benefits risk classification model. The childcare benefits scandal has illustrated how the use of discriminatory algorithms contributed to the likelihood that certain groups of people would be confronted with harsh treatment by the tax authorities. If the potential harm inflicted on the affected families in violation of their human rights had been examined prior to and during the use of the risk-scoring system, the tax authorities would have been able to identify and mitigate the risks and impacts of their policies and practices. In addition, external review of the tax authorities' algorithmic decision-making system by an independent oversight body with relevant expertise could have taken place more swiftly if a human rights impact assessment had been carried out and the oversight body had been included in that process.

<sup>144</sup> Committee on the Elimination of Racial Discrimination, General Recommendation No. 36 on preventing and combating racial profiling by law enforcement officials, 17 December 2020, UN Doc. CERD/C/GC/36, para. 60.

<sup>145</sup> Recommendation CM/REC(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems, adopted by the Committee of Ministers on 8 April 2020 at the 1373<sup>rd</sup> meeting of the Minister's Deputies, para. 5.2.

<sup>146</sup> Amnesty International, "Amnesty International Submission to the European Commission's Consultation on Artificial Intelligence", June 2020, Recommendation 3, [ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/public-consultation\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/public-consultation_en).

<sup>147</sup> Committee of Ministers to Member States, Recommendation on the human rights impacts of algorithmic systems, 8 April 2020, para. 5.2.

<sup>148</sup> Amnesty International, June 2020, Recommendation 3.

<sup>149</sup> Amnesty International, June 2020, Recommendation 5.

Amnesty International, "Race, borders and digital technologies: submission to the UN Special Rapporteur on Contemporary Forms of Racism, Xenophobia and Related Intolerance", September 2020, Recommendation 2, [ohchr.org/Documents/Issues/Racism/SR/RaceBordersDigitalTechnologies/Amnesty\\_International.pdf](https://ohchr.org/Documents/Issues/Racism/SR/RaceBordersDigitalTechnologies/Amnesty_International.pdf).  
European Digital Rights, "Recommendations for a Fundamental Rights-based Artificial Intelligence Regulation: Addressing collective harms, democratic oversight and impermissible use", 4 June 2020, Recommendation 6, [edri.org/wp-content/uploads/2020/06/AI\\_EDRiRecommendations.pdf](https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf).

The data processing of tax authorities is subject to the provisions of the General Data Protection Regulation (GDPR). General Data Protection Regulation, Article 2(1). Article 35 of the GDPR states that data controllers are required to carry out a Data Protection Impact Assessment (DPIA). However, the DPIA is not an appropriate tool for the assessment of human rights risks. The DPIA is limited in scope and focuses solely on the processing of personal data, not on the socio-economic circumstances in which the system is used. See Committee of Ministers to Member States, Recommendation on the human rights impacts of algorithmic systems, 8 April 2020, recital 5.2; European Commission Independent High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, 8 April 2019, p. 15.

**Governments must implement a mandatory and binding human rights impact assessment in public sector use of algorithmic decision-making systems, including by law enforcement authorities. This impact assessment must be carried out during the design, development, use and evaluation phases of algorithmic decision-making systems. The impact on all relevant human rights, including on social and economic rights, must be evaluated and properly addressed in the human rights impact assessment. The process should involve relevant stakeholders, including independent human rights experts, oversight bodies, individuals from potentially impacted, marginalized and/or disadvantaged communities, and technical experts.**

## 7.5 FAILING SUPERVISION: NO HUMAN RIGHTS OVERSIGHT

*Oversight on the tax authorities' use of the childcare benefits risk classification model failed, due to the fragmented nature of existing oversight mechanisms, the lack of binding human rights oversight mechanisms, and opacity of the operations of the tax authorities.*

National monitoring bodies, independent from the executive branch of government, are of great importance in contributing to meaningful human rights protections.<sup>150</sup> Strong human rights institutions at the national level should ensure that human rights are protected and advanced in a sustained manner.<sup>151</sup> International human rights law encourages the establishment of national, independent institutions to promote and monitor the implementation of human rights instruments.<sup>152</sup> States should therefore ensure, with the appropriate involvement of national parliaments, that policies and legislation comply fully with human rights, by systematically checking the compatibility of draft legislation and administrative practices.<sup>153</sup>

Independent oversight mechanisms play a crucial role in the context of algorithmic decision-making systems, as there are serious risks related to the use of algorithmic decision-making systems by states for decision-making in the public sector.<sup>154</sup> Significant expertise on all relevant human rights standards, including data protection compliance as well as data science, the design of algorithmic systems, and automated or semi-automated decision-making, is required to understand the possible dangers of algorithmic decision-making systems and to effectively keep pace with the introduction and ongoing development of algorithmic decision-making systems in the public sector.<sup>155</sup>

The practices of the Dutch tax authorities regarding fraud detection in the context of childcare benefits have been examined by various institutions. However, this oversight was fragmented and not comprehensive. As early as 2017, the National Ombudsman released a non-binding report, *No power play, but fair play*,<sup>156</sup> strongly criticizing the tax authorities and recommending that affected individuals should be provided with

<sup>150</sup> Ineke Boerefijn, "International Human Rights in National Law", in Catarina Krause and Martin Scheinin (editors), *International Protection of Human Rights: A Textbook*, 2009, pp. 587-595. Principles relating to the Status of National Institutions (The Paris Principles), 20 December 1993, Principle 3.

<sup>151</sup> United Nations, "Strengthening of the United Nations: An Agenda for Further Change", 9 September 2002, para. 50.

<sup>152</sup> Ineke Boerefijn, 2009, p. 594. See, for example, Committee on the Elimination of Racial Discrimination, General Recommendation number 17, 1993, paras. 249–250.

<sup>153</sup> Copenhagen Declaration, 12-13 April 2018, para. 16.

<sup>154</sup> Committee on the Elimination of Racial Discrimination, General Recommendation No. 36 on preventing and combating racial profiling by law enforcement officials, 17 December 2020, UN Doc. CERD/C/GC/36, para. 62. Amnesty International and Access Now, The Toronto Declaration, 16 May 2018, Articles 6 and 33, [torontodeclaration.org](http://torontodeclaration.org).

<sup>155</sup> Amnesty International, *We Sense Trouble: Automated Discrimination and Mass Surveillance in Predictive Policing in the Netherlands* (Index: EUR 35/2971/2020), 28 September 2020, p. 44, [amnesty.org/en/documents/eur35/2971/2020/en](http://amnesty.org/en/documents/eur35/2971/2020/en).

<sup>156</sup> The original title of the report in Dutch is *Geen powerplay maar fair play*.

compensation.<sup>157</sup> Partly due to the non-binding status of the report, the warnings and recommendations by the Ombudsman were ignored by the government.<sup>158</sup> In addition, despite a separate investigation into the childcare benefits risk classification model, the Ombudsman proved unable to uncover the truth about the tax authorities' use of applicants' nationality, dual citizenship or migration background in the risk classification model.<sup>159</sup> Moreover, the Netherlands Institute for Human Rights, the general supervisory body investigating human rights violations in the Netherlands, only started its investigation after publication of the report by the parliamentary investigation committee in 2020, despite having received multiple complaints from the victims of the childcare benefits risk classification model over the years.<sup>160</sup> The conclusions of the Netherlands Institute for Human Rights have not yet been published at the time of writing. However, the recommendations of the Netherlands Institute for Human Rights are not legally binding and, as has happened previously with similar documents, can be disregarded by the Dutch government.<sup>161</sup>

In contrast, the Dutch Data Protection Authority does have the power to issue binding opinions. That supervisory authority started its investigations in 2019, even though affected parents and caregivers raised concerns about the childcare benefits fraud detection policies with that body as early as 2017.<sup>162</sup>

The Dutch Data Protection Authority has stated that it was repeatedly misled by the tax authorities.<sup>163</sup> In the course of its investigations, the Dutch Data Protection Authority connected the type of data points to the type of discrimination, which is an incorrect interpretation of international human rights law. Even so, it concluded that the tax authorities processed "nationality" as a data point in the risk classification model and that this processing was unlawful and discriminatory on the basis of nationality.<sup>164</sup>

The Dutch Data Protection Authority was unable to determine whether the tax authorities had processed data points on ethnicity, such as race or skin colour, and argued that (in general) nationality should not be

<sup>157</sup> Nationale Ombudsman, *Geen powerplay maar fair play: onevenredig harde aanpak van 232 gezinnen met kinderopvangtoeslag*, 9 August 2017, p. 4, [nationaleombudsman.nl/system/files/onderzoek/Rapport%202017-095%20Geen%20powerplay%20maar%20fair%20play\\_0.pdf](https://www.nationaleombudsman.nl/system/files/onderzoek/Rapport%202017-095%20Geen%20powerplay%20maar%20fair%20play_0.pdf).

<sup>158</sup> NU.nl, "Ombudsman zegt dat overheid hem jarenlang niet hoorde in Toeslagenaffaire", 7 March 2021, [nu.nl/politiek/6120200/ombudsman-zegt-dat-overheid-hem-jarenlang-niet-hoorde-in-toeslagenaffaire.html](https://nu.nl/politiek/6120200/ombudsman-zegt-dat-overheid-hem-jarenlang-niet-hoorde-in-toeslagenaffaire.html).

<sup>159</sup> Nationale Ombudsman, "Antwoord op vragen RTL en Trouw van 12 april 2019", 17 April 2019, [rtlnieuws.nl/sites/default/files/content/documents/2019/05/19/Antwoord%20Nationale%20ombudsman%20op%20vragen%20van%20RTL%20en%20Trouw.pdf](https://rtlnieuws.nl/sites/default/files/content/documents/2019/05/19/Antwoord%20Nationale%20ombudsman%20op%20vragen%20van%20RTL%20en%20Trouw.pdf).

<sup>160</sup> Netherlands Institute for Human Rights, "Onderzoek College naar tientallen klachten over de Belastingdienst", 17 June 2021, [mensenrechten.nl/nl/nieuws/onderzoek-college-naar-tientallen-klachten-over-de-belastingdienst](https://mensenrechten.nl/nl/nieuws/onderzoek-college-naar-tientallen-klachten-over-de-belastingdienst).

<sup>161</sup> The Netherlands Institute for Human Rights has made numerous recommendations to prevent discrimination and has stated its opinion that the Dutch government is not doing enough to address the systemic nature of racism in the Netherlands. See, for example, Netherlands Institute for Human Rights, "College tegen VN: rassendiscriminatie komt op alle domeinen voor in Nederland", 17 August 2021, [mensenrechten.nl/nl/nieuws/college-tegen-vn-rassendiscriminatie-komt-op-alle-domeinen-voor-nederland](https://mensenrechten.nl/nl/nieuws/college-tegen-vn-rassendiscriminatie-komt-op-alle-domeinen-voor-nederland).

Multiple calls by the Netherlands Institute for Human Rights to implement anti-discrimination safeguards in legislative proposals have been disregarded by the Dutch government. See, for example, Netherlands Institute for Human Rights, "Intrekken nationaliteit uitreizigers op gespannen voet met gelijke behandeling", 23 July 2021, [mensenrechten.nl/nl/toegelicht/intrekken-nationaliteit-uitreizigers-op-gespannen-voet-met-gelijke-behandeling](https://mensenrechten.nl/nl/toegelicht/intrekken-nationaliteit-uitreizigers-op-gespannen-voet-met-gelijke-behandeling).  
Netherlands Institute for Human Rights, "Gelijke behandeling en rechtsbescherming in de knel bij Wet gegevensverwerking door samenwerkingsverbanden", 25 June 2021, [mensenrechten.nl/nl/nieuws/gelijke-behandeling-en-rechtsbescherming-de-knel-bij-wet-gegevensverwerking-door](https://mensenrechten.nl/nl/nieuws/gelijke-behandeling-en-rechtsbescherming-de-knel-bij-wet-gegevensverwerking-door).

In 2020, the UN Human Rights Committee found that the Netherlands had violated the rights of a child by registering them as 'nationality unknown' and not recognizing them as stateless. The Netherlands Institute for Human Rights had previously called on the Netherlands to address the inadequate protection of children's rights, to no avail. See Netherlands Institute for Human Rights, "Nederland moet betere bescherming bieden aan kinderen zonder nationaliteit", 6 January 2021, [mensenrechten.nl/nl/nieuws/nederland-moet-betere-bescherming-bieden-aan-kinderen-zonder-nationaliteit](https://mensenrechten.nl/nl/nieuws/nederland-moet-betere-bescherming-bieden-aan-kinderen-zonder-nationaliteit).

<sup>162</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 6.

<sup>163</sup> Dutch Data Protection Authority, "Presentatie onderzoeksrapport 'De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag' door Aleid Wolfsen op 17 juli 2020", 17 July 2020, [autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/toespraak\\_aleid\\_wolfesen\\_onderzoek\\_kinderopvangtoeslag.pdf](https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/toespraak_aleid_wolfesen_onderzoek_kinderopvangtoeslag.pdf).  
The tax authorities were uncooperative and untransparent to the parliamentary investigation committee as well. Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 8.

<sup>164</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 50.

considered a direct proxy for ethnicity or race data.<sup>165</sup> It continued to investigate whether, under the data protection criteria, the processing of nationality data by the tax authorities should be considered the processing of sensitive data, specifically looking at whether the tax authorities had the explicit purpose to discriminate based on ethnicity or if discrimination based on ethnicity could be “reasonably foreseen”.<sup>166</sup> However, under international law, the existence of prejudice, or an intention to discriminate, are not relevant to determine whether the legal criteria for discrimination has been satisfied. It is the outcome of the treatment that determines whether there is discrimination (see Section 6.1). By adding data protection criteria to the mix, the Dutch Data Protection Authority failed to answer a crucial question: does the processing of nationality data lead to differential treatment on the base of ethnicity, and is this difference objectively and reasonably justified?

The Dutch Data Protection Authority’s approach for the assessment of discrimination underscores the need for a supervisory body with binding powers that focuses on the human rights impacts of algorithmic decision-making systems, alongside the data protection focus that is offered by data protection supervisory authorities. It demonstrates that human rights oversight mechanisms should include investigation of the objective and rationale of an algorithmic decision-making system and why certain risk factors, such as nationality or even seemingly neutral factors, are used in the system, as well as the direct and/or indirect discriminatory effect of the algorithmic system.

**Governments should establish comprehensive and independent human rights oversight mechanisms over the use of algorithmic decision-making systems, which includes oversight over civil and political as well as social and economic rights, to strengthen accountability mechanisms and increase human rights protection. An oversight body should be granted the mandate, powers, and capacity to oversee human rights protection in the use of algorithmic decision-making systems in the public sector, issue guidance, and hold designers, developers, and users to account in binding decisions.**

**SUCH AN OVERSIGHT BODY SHOULD AT THE VERY LEAST:**

- Be granted the power to oversee how government bodies respect, protect and fulfil all human rights, including economic, social and cultural rights, the right to good governance, and the right to equality and non-discrimination;
- Have the required technical and legal knowledge, as well as sufficient capacity and resources to carry out its duties;
- Actively and proactively investigate human rights violations and risks related to the use of algorithmic decision-making systems in the public sector;
- Be able to receive and handle complaints related to the impact of algorithmic decision-making systems from affected individuals, whistle-blowers and civil society;
- Have the power to issue administrative sanctions on designers, developers or users of algorithmic decision-making systems that result in human rights violations or harm, or if there is a high and unmitigated risk of such impacts;
- Have access to the documentation, human rights impact assessments, training data, data categories, and algorithms in order to examine the algorithmic decision-making systems and their outcomes;
- Regularly consult with relevant stakeholders, such as human rights and technical experts, and in particular impacted, marginalized and/or disadvantaged communities.
- Have a mandate that does not interfere with any individual’s right to seek a remedy in a court (see Section 7.7).

<sup>165</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, p. 35.

<sup>166</sup> Dutch Data Protection Authority, Report (previously cited), 16 July 2020, pp. 35 and 48.

## 7.6 DISCRIMINATION IN PLAIN SIGHT: NO BAN ON RACIAL PROFILING IN THE NETHERLANDS

*While the Dutch government publicly opposes racial profiling, it continues to allow the use of ethnicity and other prohibited grounds of discrimination as a risk factor for suspicion and as a basis for decision-making in law enforcement.*

In addition to the obligation to refrain from discriminatory measures, states also have a *positive obligation* to prevent, stop or punish discrimination.<sup>167</sup> This obligation includes, when necessary, the adoption of legislation that prohibits certain actions that have a high risk of resulting in discrimination.<sup>168</sup>

In a debate on the childcare benefits scandal in January 2021, Mark Rutte, Prime Minister of the Netherlands, refused to stop the use of nationality in risk-profiling and risk-scoring.<sup>169</sup> This position has also been confirmed in writing by the Cabinet in response to various motions and proposals to prohibit the use of ethnicity and nationality in risk profiles used for law enforcement.<sup>170</sup> Earlier in 2020, the Dutch Ministry of Justice and Security held that the use of nationality in risk profiles is permitted, as long as it is not the sole factor in a risk assessment.<sup>171</sup>

However, this argument is extremely problematic. The use of nationality, even among other factors, in a risk-scoring system deployed by law enforcement leads to different outcomes and different treatment depending on the nationality of certain individuals.<sup>172</sup> Section 6.1 details how this leads to racial profiling and disproportionately affects people from ethnic minority backgrounds.

The discriminatory practice of using data on nationality and ethnicity, or proxies thereof, in risk-profiling for law enforcement purposes in search of potential criminal suspects continues in the Netherlands, and the Dutch government continues to fail in its positive obligation to avoid discrimination and in its obligation to take effective measures to end racial profiling.<sup>173</sup> In June 2021, the Dutch Human Rights Institute issued

<sup>167</sup> ECtHR, *Pla and Puncernau v. Andorra*, Application 69498/01, 13 July 2004, para. 62, [hudoc.echr.coe.int/eng?i=001-61900](https://hudoc.echr.coe.int/eng?i=001-61900). ECtHR, *Horváth and Kiss v. Hungary*, Application 11146/11, 29 January 2013, paras. 116, 119 and 127, [hudoc.echr.coe.int/eng?i=001-116124](https://hudoc.echr.coe.int/eng?i=001-116124).

<sup>168</sup> Committee on the Elimination of Racial Discrimination, General Recommendation 36, 24 November 2020, UN Doc. CERD/C/GC/36, para. 27.

<sup>169</sup> Netherlands House of Representatives, “Debat over de verklaring van de minister-president en over het verslag van de ondervragingscommissie Kinderopvangtoeslag”, 19 January 2021, [tweedekamer.nl/debat\\_en\\_vergadering/plenaire\\_vergaderingen/details/activiteit?id=2021A00393](https://tweedekamer.nl/debat_en_vergadering/plenaire_vergaderingen/details/activiteit?id=2021A00393). In this debate regarding the report of the parliamentary investigative committee, the Prime Minister of the Netherlands stated that he was “in principle” willing to exclude nationality, ethnicity and place of birth as risk indicators, “unless there is an objective justification for such data processing”. It should be noted that the state holds the same position in a lawsuit about ethnic profiling by the Dutch border police, see The District Court of The Hague, 15 June 2021, C/09/589067 / HA ZA 2020-35, [p1pnljcm.nl/wp-content/uploads/2021/06/20210615-Pleadings-English-007A.pdf](https://p1pnljcm.nl/wp-content/uploads/2021/06/20210615-Pleadings-English-007A.pdf).

<sup>170</sup> Ministry of the Interior and Kingdom Relations, “Waardering motie Azarkan/Van Weyenberg over een verbod op gebruik van nationaliteit of etniciteit bij risicoprofilering”, 9 February 2021, [rijksoverheid.nl/documenten/kamerstukken/2021/02/09/kamerbrief-reactie-op-motie-verbod-gebruik-nationaliteit-of-etniciteit-bij-risicoprofilering](https://rijksoverheid.nl/documenten/kamerstukken/2021/02/09/kamerbrief-reactie-op-motie-verbod-gebruik-nationaliteit-of-etniciteit-bij-risicoprofilering).

<sup>171</sup> Ministry of Justice and Security, “Reactie op rapport Amnesty International over massasurveillance”, 11 December 2020, [rijksoverheid.nl/documenten/kamerstukken/2020/12/11/tk-reactie-op-rapport-amnesty-international-over-massasurveillance](https://rijksoverheid.nl/documenten/kamerstukken/2020/12/11/tk-reactie-op-rapport-amnesty-international-over-massasurveillance).

<sup>172</sup> As outlined in Section 6.1, there are very limited circumstances in which the use of personal characteristics that relate to nationality and/or ethnicity can be considered legitimate for law enforcement purposes. Risk-scoring is, however, fundamentally different from targeted investigation because there is no individualized suspicion under investigation that might justify the use of the criteria of nationality and/or ethnicity in combination with other criteria which can be objectively justified.

<sup>173</sup> A civil society coalition, which includes Amnesty International, has initiated a court case against the Dutch border police (the Royal Netherlands Marechaussee) for racial profiling. The coalition plans to go into appeal against the ruling in first instance, because it was silent on the use of nationality and ethnicity in risk profiles, and allowed racial profiling in enforcement decisions. See Amnesty International the Netherlands, *Dutch court allows racial profiling*, 22 September 2021, [amnesty.nl/content/uploads/2021/09/Dutch-Court-allows-racial-profiling\\_September-22\\_2021\\_final.pdf?x88229](https://amnesty.nl/content/uploads/2021/09/Dutch-Court-allows-racial-profiling_September-22_2021_final.pdf?x88229) and [rechtspraak.nl/uitspraak?id=ECLI:NL:RBDHA:2021:10283](https://rechtspraak.nl/uitspraak?id=ECLI:NL:RBDHA:2021:10283). For more information about this case, see Amnesty International the Netherlands, “Rechtszaak tegen Marechaussee vanwege etnisch profileren”, [amnesty.nl/rechtszaak-tegen-marechaussee](https://amnesty.nl/rechtszaak-tegen-marechaussee).

a publication on racial profiling and the use of race and ethnicity in risk profiles used for fraud detection involving, for example, social benefits, stop-and-search activity by law enforcement, and border controls.<sup>174</sup> The publication describes the public debate surrounding this issue and outlines the relevant human rights norms in relation to racial profiling and the use of race and ethnicity in risk profiles used for fraud detection, but does not take a clear-cut position. The Dutch government should close the door on any use of race or ethnicity as criteria in risk-profiling and assessments of people against whom there is no individualized suspicion of any wrongdoing.

**Governments must establish a clear, unambiguous and legally binding ban on the use of data regarding nationality and ethnicity, or proxies thereof, in risk scoring for law enforcement purposes in search of potential crime or fraud suspects.**

## 7.7 KAFKA'S CASTLE: A LACK OF EFFECTIVE REMEDIES AND REDRESS

*The lack of transparency on and accountability of the tax authorities drove the victims of the childcare benefits fraud system into Kafkaesque<sup>175</sup> legal procedures, with no effective remedies open to the parents and caregivers that were discriminated against with the risk classification model. Currently, remedial action taken by the Dutch government for parents and caregivers affected by the childcare benefits scandal excludes redress for discrimination caused by the risk classification model.*

When human rights violations occur, international law requires that individuals are guaranteed the right to an effective remedy and the right to adequate redress.<sup>176</sup> These rights include that the victim is granted access to relevant information concerning violations and reparation mechanisms. Effective remedies – or measures to repair the harm caused to victims of human rights violations – can take various forms and depend on the nature of the right violated, the harm suffered and the wishes of those affected.<sup>177</sup> Compensation is a form of reparation to offset damages sustained as a result of an infringement of human rights.<sup>178</sup>

Victims of human rights violations frequently face significant challenges when seeking access to justice. The UN Special Rapporteur on Freedom of Expression has highlighted how algorithmic decision-making systems in general often interfere with the right to remedy.<sup>179</sup> There is an inherent challenge around access to information, as “individuals are not aware of the scope, extent or even existence of algorithmic decision-making systems that are affecting their rights”. This opacity is exacerbated by the fact that algorithms are constantly adapting and changing, such that even the designers of the system may not be able to explain

<sup>174</sup> Netherlands Institute for Human Rights, *Debat over etnisch profileren vraagt om meer juridische duidelijkheid*, 9 June 2021, [publicaties.mensenrechten.nl/file/a261a614-6d6e-4d1d-be8c-0c88fd43b954.pdf](https://publicaties.mensenrechten.nl/file/a261a614-6d6e-4d1d-be8c-0c88fd43b954.pdf).

<sup>175</sup> Kafkaesque refers to themes from the work of Franz Kafka. For example, his 1926 novel *The Castle* is often understood to be about alienation, unresponsive bureaucracy, and the frustrations of trying to conduct business with non-transparent, seemingly arbitrary controlling systems. Similarly, his 1925 novel *The Trial* describes how a man is arrested and prosecuted by a remote, inaccessible authority, with the nature of his crime never revealed.

<sup>176</sup> Under international human rights standards, the notion of access to justice is enshrined in the European Convention for the Protection of Human Rights and Fundamental Freedoms, Article 6 and 13; EU Charter of Fundamental Rights, Article 47. These rights are also provided for in other international instruments, such as UN International Covenant on Civil and Political Rights, Articles 2(3) and 14; UN Universal Declaration of Human Rights Articles 8 and 10; International Covenant on Economic, Social and Cultural Rights, Article 2; International Convention on the Elimination of All Forms of Racial Discrimination, Article 6.

<sup>177</sup> European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European law relating to access to justice*, January 2016, p. 101, [fra.europa.eu/en/publication/2016/handbook-european-law-relating-access-justice](https://fra.europa.eu/en/publication/2016/handbook-european-law-relating-access-justice).

<sup>178</sup> European Union Agency for Fundamental Rights and Council of Europe, January 2016, p. 102.

<sup>179</sup> UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 29 August 2018, UN Doc. A/73/348, para. 40.

how they reached their outcomes, or do so poorly.<sup>180</sup> The onus is on states to make algorithmic decision-making systems visible, to allow outputs or impacts to be queried and appealed, and to create accessible and practical routes for remedy and redress when human rights are negatively impacted.<sup>181</sup> The key attributes of good governance play an important role in the realization of human rights in the data-driven society. States should ensure that all instances of algorithmic bias are duly investigated and that sanctions are imposed.<sup>182</sup>

The childcare benefits scandal included many obstacles for the victims to access justice and obtain effective remedies after suffering adverse human rights impacts. In part, this is due to inherent challenges that algorithmic decision-making systems pose to obtaining access to remedy. Parents and caregivers who were identified by the tax authorities as fraudsters were for years given no answers to questions about what they had done wrong.<sup>183</sup> Requests for information to inspect files were often refused by the tax authorities.<sup>184</sup> Complaints from the parents and caregivers filed with the tax authorities were handled two years after the date of complaint.<sup>185</sup> The opacity of the risk-scoring system and other practices resulted in affected parents and caregivers being denied accessible and practical routes for remedy and redress. Despite years of suspicion, affected parents and caregivers had no idea that they were being racially profiled. Even after the news received widespread national attention, regulators did not address the question of legal accountability.

Years after the harm was done, remedial measures for the affected parents and caregivers have finally been taken. The harms for which the parents and caregivers are eligible for compensation exclude, however, discrimination by the tax authorities on the basis of nationality, ethnicity or social origin as a result of using the risk classification model.<sup>186</sup> It is only in circumstances where additional discrimination can be observed in the further handling of a case by civil servants that a parent or caregiver may be eligible for compensation on the basis of discrimination.<sup>187</sup> There are no effective remedies or redress for racial profiling and other forms of discrimination stemming from the use of the algorithmic risk classification model.

States must provide meaningful accountability, effective remedy and redress for human rights harms linked to the use of algorithmic decision-making systems. This may include, for example, creating clear, independent, accessible processes for redress and designating roles in the public sector that are responsible for the timely remedy of human rights harms, subject to accessible and effective appeal and judicial review.

---

<sup>180</sup> *AI Now, AI Now 2017 Report*, 2017, p. 30, [ainowinstitute.org/AI\\_Now\\_2017\\_Report.pdf](https://ainowinstitute.org/AI_Now_2017_Report.pdf).

<sup>181</sup> Amnesty International (previously cited), June 2020.

<sup>182</sup> Committee on the Elimination of Racial Discrimination, General Recommendation No. 36 on preventing and combating racial profiling by law enforcement officials, 17 December 2020, UN Doc. CERD/C/GC/36, para. 65.

<sup>183</sup> NOS, “Advocaat: toeslagenaffaire doet denken aan Kafka, liep tegen muur aan”, 16 November 2020, [nos.nl/artikel/2356771-advocaat-toeslagenaffaire-doet-denken-aan-kafka-liep-tegen-muur-aan](https://nos.nl/artikel/2356771-advocaat-toeslagenaffaire-doet-denken-aan-kafka-liep-tegen-muur-aan).

<sup>184</sup> Jan Kleinnijenhuis, “Gedupeerde ouders willen inzage in hun dossier, maar de Belastingdienst voelt daar weinig voor”, *Trouw*, 25 August 2019, [trouw.nl/nieuws/gedupeerde-ouders-willen-inzage-in-hun-dossier-maar-de-belastingdienst-voelt-daar-weinig-voor-b6771696](https://trouw.nl/nieuws/gedupeerde-ouders-willen-inzage-in-hun-dossier-maar-de-belastingdienst-voelt-daar-weinig-voor-b6771696).

<sup>185</sup> Netherlands House of Representatives, Report (previously cited), 17 December 2020, p. 77.

<sup>186</sup> Prime Minister of the Netherlands, Minister of General Affairs, Letter to Parliament in response to the report “Ongekend Onrecht”, 15 January 2021, pp. 2-3, [rijksoverheid.nl/documenten/kamerstukken/2021/01/15/kamerbrief-met-reactie-kabinet-op-rapport-ongekend-onrecht](https://rijksoverheid.nl/documenten/kamerstukken/2021/01/15/kamerbrief-met-reactie-kabinet-op-rapport-ongekend-onrecht).

<sup>187</sup> Prime Minister of the Netherlands, Minister of General Affairs, 15 January 2021, p. 16.

# 8. RECOMMENDATIONS

## TO ALL STATES, INCLUDING THE NETHERLANDS:

### 1. Protect human rights in the use of algorithmic decision-making systems

Governments must put in place a framework that a) prevents human rights violations in relation to the use of algorithmic decision-making systems from taking place, b) establishes monitoring and oversight mechanisms as safeguards, c) holds those responsible for violations to account, and d) provides effective remedy to individuals and groups whose rights have been violated.

### 2. No black box systems in the public sector

Governments must refrain from using black box systems in high-risk contexts, such as in the context of social security fraud detection.

### 3. Restrictions of self-learning algorithms for decision-making that has significant impacts

Governments must ban the use of self-learning algorithms in the public sector at the point of deployment for a) decision-making that results in legal effect, b) decision-making and actions that affect the rights and freedoms of individuals, and c) decision-making and actions that have a major impact on society.

### 4. Public registry for algorithmic decision-making systems

Governments should create the maximum possible transparency regarding public sector use of algorithmic decision-making systems by creating a public registry which includes detailed and comprehensive information of public sector use of algorithmic decision-making systems.

### 5. Transparency of semi-automated decision-making systems

Governments should be fully transparent and provide meaningful information to affected individuals about the underlying logic, importance and expected consequences of decisions, even if they are not *fully* automated, regardless of the level of human involvement in the decision-making process.

### 6. Human rights impact assessment

Governments must implement a mandatory and binding human rights impact assessment in public sector use of algorithmic decision-making systems, including by law enforcement authorities. This impact assessment must be carried out during the design, development, use and evaluation phases of algorithmic decision-making systems. The impact on all relevant human rights, including on social and economic rights, must be assessed and properly addressed in the human rights impact assessment. The process should involve relevant stakeholders, including independent human rights experts, oversight bodies, individuals from potentially impacted, marginalized and/or disadvantaged communities, and technical experts.

## **7. Comprehensive human rights supervision**

Governments should establish comprehensive and independent human rights oversight mechanisms over the use of algorithmic decision-making systems, which includes oversight over civil and political as well as social and economic rights to strengthen accountability mechanisms and increase human rights protection. An oversight body should be granted the mandate, powers, and capacity to oversee human rights protection in the use of algorithmic decision-making systems in the public sector, issue guidance, and hold designers, developers, and users to account in binding decisions.

## **8. No profiling on the basis of race or ethnicity**

Governments must establish a clear, unambiguous and legally binding ban on the use of data regarding nationality and ethnicity, or proxies thereof, in risk-scoring in search of potential perpetrators, or crime or fraud suspects.

## **9. Accountability, effective remedy and redress**

Governments must provide meaningful accountability, effective remedy and redress for human rights harms linked to the use of algorithmic decision-making systems. This may include, for example, creating clear, independent, accessible processes for redress and designating roles in the public sector that are responsible for the timely remedy of human rights harms, subject to accessible and effective appeal and judicial review.

# **TO THE DUTCH GOVERNMENT:**

## **1. Provide adequate compensation to the victims of discrimination through the use of the risk classification model**

In addition to the compensation scheme already established for victims of other harms inflicted during the childcare benefits scandal, the Dutch government must provide compensation for discrimination on the basis of nationality, ethnicity, social origin and racial profiling to the parents and caregivers that were selected by the risk classification model partly on the basis of their non-Dutch nationality.

The government must provide adequate compensation in swift and effective procedures for all harms identified.

## **2. Ratify Convention 108+**

The Dutch government should ratify and implement Convention 108+.

**AMNESTY INTERNATIONAL  
IS A GLOBAL MOVEMENT  
FOR HUMAN RIGHTS.  
WHEN INJUSTICE HAPPENS  
TO ONE PERSON, IT  
MATTERS TO US ALL.**

## CONTACT US



[info@amnesty.org](mailto:info@amnesty.org)



+44 (0)20 7431 5500

## JOIN THE CONVERSATION



[www.facebook.com/AmnestyGlobal](https://www.facebook.com/AmnestyGlobal)



[@Amnesty](https://twitter.com/Amnesty)

# XENOPHOBIC MACHINES

## DISCRIMINATION THROUGH UNREGULATED USE OF ALGORITHMS IN THE DUTCH CHILDCARE BENEFITS SCANDAL

Social security enforcement agencies worldwide are increasingly automating their processes in the hope of detecting fraud. The Netherlands is at the forefront of this development. The Dutch tax authorities adopted an algorithmic decision-making system to create risk profiles of individuals applying for childcare benefits in order to detect inaccurate and potentially fraudulent applications at an early stage. Nationality was one of the risk factors used by the tax authorities to assess the risk of inaccuracy and/or fraud in the applications submitted. This report illustrates how the use of individuals' nationality resulted in discrimination based on nationality and ethnicity, as well as racial profiling.

Amnesty International calls on governments to put in place frameworks that a) prevent human rights violations in relation to the use of algorithmic decision-making systems from taking place, b) establish monitoring and oversight mechanisms, c) hold those responsible to account, and d) provide effective remedy to individuals and groups whose rights have been violated. The Dutch government must provide adequate, effective and swift compensation to victims of discrimination resulting from the use of their nationality as a risk factor in the childcare benefits risk-scoring system.

INDEX: EUR 35/4686/2021  
OCTOBER 2021  
ORIGINAL LANGUAGE: ENGLISH

[amnesty.org](https://www.amnesty.org)

**AMNESTY**  
INTERNATIONAL

