12. Motherhood in the Datafied Welfare State: Investigating the Gendered and Racialized Enactment of Citizenship in Dutch Algorithmic Governance

Gerwin van Schie, Laura Candidatu, and Diletta Huyskes

Abstract: In 2023, Rotterdam discontinued an invasive, biased welfare fraud risk-scoring algorithm after an investigative report by Lighthouse Reports, which exposed its racial and gender biases, disproportionately affecting migrant mothers in deprived areas. This chapter argues that such biases could have been identified before implementation by scrutinizing the categories embedded in the algorithm and contextualizing them within the history of the Dutch welfare system. Using a genealogical approach, we trace how norms about race and gender became embedded in welfare practices. A category analysis shows how these biases shaped the algorithm's indicators. Drawing on critical data studies and feminist theories on migrant motherhood and racialized citizenship, we show how discriminatory ideas about the "ideal" welfare recipient predate the algorithm, contributing to discussions about equality in datafied welfare governance.

Keywords: welfare surveillance, algorithmic auditing, intersectionality, discriminatory algorithms, feminist critical data studies

Introduction

On March 6, 2023, the investigative journalism platform Lighthouse Reports published a critical article on *Wired.com* addressing how the Dutch city of Rotterdam had been using an invasive and biased welfare fraud risk-scoring

Van Dijck, José, Karin van Es, Anne Helmond, and Fernando van der Vlist, eds. *Governing the Digital Society: Platforms, Artificial Intelligence, and Public Values*. Amsterdam: Amsterdam University Press, 2025
DOI 10.5117/9789048562718_CH12

algorithm since 2018. Using information obtained through a request to the administration, the investigative team at Lighthouse Reports was able to test the algorithmic model used by the municipality of Rotterdam on anonymized data of inhabitants (Constantaris et al. 2023). This allowed them to conduct realistic "tests" to check how the model would perform under specific scenarios considering different risk indicators (Casimir-Braun et al. 2023). From these tests, they concluded that the archetype of a non-Dutch mother struggling financially is among the most at-risk profiles, and it gets even worse when that person is young and with limited proficiency in Dutch. After an internal investigation by the city of Rotterdam, algorithmic risk scoring as a work practice was discontinued.

While many journalists and scholars have treated this case as: (1) exemplary of new problems associated with the datafication and algorithmization of government bureaucracy and (2) a deviation from ideals of equality and justice, we find that it signifies a continuation of gendered and racialized policing practices with a much longer history. We will make this argument by combining a genealogical lens (Tamboukou 1999, 202–3), aimed at tracing how norms about race and gender came to be reproduced in Dutch public institutions' welfare distribution processes, with a category analysis (Yanow et al. 2016), aimed at revealing the norms embedded in the indicators of the welfare fraud risk-scoring algorithm. Where the genealogy is based on existing literature on Dutch welfare, the category analysis entails a critical study of the indicators present in the risk-scoring calculation as detailed in a public letter by Richard Moti (2021), the alderman of the Work and Income Department of the municipality of Rotterdam at the time. By linking past welfare discourses and contemporary algorithmic systems, we argue that contemporary bureaucratic practices continue the enforcement of sexist and racist ideas about the ideal beneficiary of state welfare that were already present in Dutch welfare discourse. The social positioning of migrant mothers is thus naturalized within the algorithm, which ultimately questions their subjectivity as citizens and their rights as citizens. With this critical analysis of the categories within a data assemblage, we advocate for auditing approaches that take into account both the historical contexts and the prevailing discursive practices shaping and reproducing institutional processes. This way, the discriminatory potential of algorithmic systems could be mitigated before systems are implemented, instead of stopped after marginalized people (who were marginalized to begin with) have been victimized.

In this chapter, we will draw on critical data studies scholarship (Dencik and Kaun 2020; Van Zoonen 2020) as well as feminist theories on migrant motherhood and racialized citizenship (Schields 2023; Waaldijk 2007). Welfare distribution has an intimate historical connection with processes of government surveillance of underprivileged people (Kohler-Hausmann 2007). Given this intimate history, welfare distribution and surveillance have proven to be a prominent site for the implementation of data systems and algorithms in public governance (Dencik 2022; Dencik and Kaun 2020; Choroszewicz and Mäihäniemi 2020; Mann 2020). A recent literature review of 190 articles on datafied public service and social equity calls for more attention to "technical, socio-technical, and systemic mechanisms that are responsible for linking data-driven public service provision to equity" (Ruijer et al. 2023, 326–28). With this chapter, we aim to contribute to this line of investigation by explicitly linking past welfare regimes and their focus on women in general, and migrant women in particular, with contemporary algorithmic bureaucratic practices.

Governing the Dutch datafied welfare state

The implementation of the Rotterdam welfare fraud risk-scoring algorithm is part of a larger development in which governments at all levels are increasingly relying on computational systems and algorithms for the execution of bureaucratic processes. The assumption is that these systems will improve service and work "more fairly without human interference and flaws" (Dencik and Kaun 2020, 2). In the case of the Netherlands, the rapid implementation of these relatively new systems in state bureaucracy came with multiple scandals involving discriminatory algorithmic systems and work practices (Peeters and Widlak 2023; Van Bekkum and Zuiderveen Borgesius 2021). A recurring theme through most of these scandals is that the systems generally disadvantaged Dutch citizens of color and/or with a migration background (see table 12.1 for an overview). This gives the impression of deeply ingrained structures of institutional racism in Dutch governance. Importantly, we consider these cases of discriminatory policy not as the result of the biases of individual policymakers, civil servants, programmers, or other people working on government data systems and algorithms, but rather as the result of widely shared cultural and organizational values and beliefs regarding deserving and undeserving citizens.

Table 12.1. Partial List of Known Discriminatory Algorithmic Systems Implemented by Dutch Governmental Organizations

Organization	System name or type	Years in service	Type of bias
Dutch Ministry of the Interior and Kingdom Relations	Leefbaarometer (Liveability Barometer) version 1 and 2	2007–22	A map that used information about migration backgrounds of inhabitants as predictors for the liveability of streets and neighborhoods. This information was used as justification for banning specific people from social housing in "problem areas" through the Rotterdam Act (Van Gent et al. 2018; Van Schie et al. 2020). The current version of this system no longer contains these indicators.
Immigratie- en Naturalisatiedienst (Immigration and Naturalisation Service, IND)	Risk-profiling algorithm	2011–22	Used birthplace and nationalities of company owners as predictors for tax fraud and illegal immigration (Van der Woude and Davidson 2022).
Dienst Uitvoering Onderwijs (Education Implementation Service, DUO)	Student bursary and loan fraud risk-scoring algorithm	2011–23	In 2023, a group of lawyers reported that out of their clients, 97 percent of 397 students accused of fraud had a migration background. The indicators used in the system are still unknown as of September 2023 (Heilbron and Kootstra 2023).
Dutch Tax and Customs Administration	Fraude Signalering Voorziening (FSV)	2013–20	In the Netherlands, 26,000 people were wrongfully accused of tax fraud. People with a migration background were targeted at a disproportionate rate. The system contained data points about people having a "non-Western appearance," nationality, and donations to mosques (Peeters and Widlak 2023; PwC 2022).
Employee Insurance Agency (UWV), the Dutch Social Insurance Bank (SVB), and the Dutch Tax and Customs Administration	System Risk Indication (SyRI)	2014–20	Only used in specific neighborhoods, often with a relatively large population of people with a migration background. Used criteria such as migration background and household composition to assess risk (Van Bekkum and Zuiderveen Borgesius 2021).
Dutch Police	Crime Anticipation System (CAS)	2015–present	Used the factor "percentage of people with a non-Western migration background" in its location-based risk-scoring algorithm during the testing phase. This marker was removed before its national rollout in 2017. However, the model is still based on past reports, which are heavily influenced by human biases (Van Schie and Oosterloo 2020). CAS is still in service as of 2023.

Organization	System name or type	Years in service	Type of bias
Municipality of Rotterdam	Welfare fraud risk-scoring algorithm	2018–21	Used race and ethnicity–related indicators such as language course enrollment and neighborhood, and several indicators related to motherhood, household composition, and partner history. The model is also based on past known incidents, which are heavily influenced by human biases (Aung et al. 2021).
Dutch Ministry of Foreign Affairs	Visa application risk-scoring algorithm	2015–22	Used nationality and gender as risk indicators for visa applications (Maleeyakul et al. 2023).

The increasing awareness of the risks of datafication and algorithmization has led to a growing public and academic debate on the idea of algorithmic accountability as a principle (Wieringa 2020) and various methods of algorithmic auditing as a means to achieve such accountability (Metaxa et al. 2021; Raji et al. 2020; Vecchione et al. 2021). Additionally, public administration has seen a proliferation of ethics tools and guidelines produced by various governmental actors, academics, and non-profit organizations in an attempt to more structurally embed ethical values and principles in the design process of algorithms (Franzke 2022). However, as noted by Franzke, almost no guidelines "provide indication of having reflected upon the fact that what one might consider to be 'the good' or 'the right' is strongly shaped by context, interests, circumstances and (implicit) ethical framework" (2022, 6). Utrecht University's Data School has also created two tools that have become popular in governmental organizations: the Data Ethics Decision Aid (DEDA), which aids civil servants in making ethical decisions about data and algorithms in new and ongoing projects (Franzke et al. 2021; Siffels et al. 2022), and the Fundamental Rights and Algorithms Impact Assessment (FRAIA), which aims to help civil servants protect the human rights of Dutch citizens in projects or bureaucratic practices that involve algorithms (Gerards et al. 2022). While these tools help in making explicit the laws and values that apply to a particular data system or algorithm aimed to be used in public governance, little attention is paid to historical power relations in terms of race, class, and gender that are already present in particular policy domains. Since such power relations are often normalized within policy contexts, it is hard for civil servants to recognize them as power relations and account for the perspective of marginalized citizens. In the next section, we use the Rotterdam welfare fraud risk-scoring algorithm and its historical context as an example to show how contemporary algorithmic governance

of welfare did not emerge in a vacuum but is a continuation of historical forms of welfare governance and their associated power structures.

Gender and race in neoliberal welfare reforms

The algorithmization of the fraud detection system and its racialized and gendered biases, we argue, is part of a policy legacy favoring neoliberal ideas of active citizenship, as well as a static and racialized model of citizenship. The Rotterdam algorithm's discriminatory outcomes are not just a result of algorithmic haphazardness or human bias. They are also influenced by the neoliberal political discourses and policy shifts that have been part of the Dutch welfare system's retrenchment since the 1970s.

A brief history of welfare reforms in the Netherlands

In the 1970s and early 1980s, in the context of Western democracies and the two major oil crises, the post-war interventionist approach to market regulation transitioned toward a neoliberal stance favoring deregulation, privatization, and welfare-state retrenchment (MacLeavy 2016, 252; Oudenampsen 2020). Previous approaches to the social aspects of citizenship rights were aimed at improving the social conditions for marginalized citizens and ensuring their equal footing in society. Welfare retrenchment policies, however, adopted a so-called "workfare" approach in which social benefits and rights became linked to individuals' efforts to (re)integrate into the labor market. This model emphasized people's personal responsibility for their integration into the labor market and their overall marginalization. Individuals who could not comply with the requirements for work reintegration were often portrayed as exploiting the system or potentially engaging in fraudulent behavior. In countries such as the US, the UK, and the Netherlands, various measures were implemented to reduce the so-called "culture of dependency" and identify fraudulent exploitations of the welfare system (MacLeavy 2016, 254). In the context of the Netherlands, Van Gerven (2019, 387) identifies three main welfare reforms that contributed to the larger neoliberal shift and the current Dutch workfarist governance model: decentralization, risk differentiation, and increased citizen participation. These policy strategies aimed to reduce the state's burden to support marginalized citizens. Consequently, the role of municipalities in providing social assistance increased, while, at the same time, following government advice to reduce the number of welfare beneficiaries. Furthermore, the primacy of "everyone is responsible for their own welfare"

mantra made labor market insiders the main "deserving" participants in the larger state solidarity network (ibid., 401). In this model, the terms and conditions for participating in state networks of solidarity are based on active citizenship, linking eligibility for state support to one's capacity for economic productivity.

The "failure of multiculturalism" and the migrant Other

As a backdrop to these economic concerns and welfare reforms, there were Dutch public debates around the so-called "failure of multiculturalism," which prompted subsequent policy changes. These debates were symptomatic of a political panic around the cultural difference and belonging of migrant minorities. The gendered and racial dimensions of citizenship come particularly to the fore in this context. First, after the initial multicultural policy approach centered on the preservation of minority cultures in the 1960s and 1970s, the preoccupation with their socioeconomic and cultural integration became stronger. Different subsequent policies promoted migrant people's participation in education and labor as a means for greater integration (Entzinger 2003, 70–72). Second, in the post-9/11 context, these reforms reproduced assumptions about gender, religious, ethnicity and race differences between minority groups. The premise was that Dutch laws and norms were in opposition with norms regulating migrant (Muslim) men's masculinity, while migrant (Muslim) women were seen as casualties of their own cultural norms (Prins and Saharso 2008, 368). Gender, ethnicity, and (racialized) religion thus became important aspects for distinguishing between the ideal national subject of the "imagined community" (Anderson 1986)—the "hard-working" (Mepschen 2012) and emancipated white Dutch citizen, on the one hand, and the non-active, non-integrated, and non-emancipated migrant Other, on the other (Gorashi and Vieten 2012, 730). These racialized and gendered understandings of national belonging continue to shape the regulation of welfare benefits. As the case of the welfare fraud risk-scoring algorithm used by the city of Rotterdam will show, they are further reproduced in institutionalized norms and bureaucratic practices that regulate citizenship and belonging to networks of solidarity.

Race, gender, and non-normative family formations

To further historicize and contextualize the specific bias of the algorithm towards single migrant mothers we now take a feminist and intersectional lens. In this, we are particularly interested in showing how race and gender together contribute to discriminatory effects in the distribution of social

benefits. More specifically, we explore the relation between the algorithm's discriminatory effects towards single migrant mothers and the racist trope of the "welfare queen" used in political discourses about welfare retrenchment.

Gender has consistently influenced how state-mediated social rights are distributed. One way in which this gendering plays out, as pointed out by various feminist scholars, refers to welfare states' initial disciplining of women through specific care arrangements centered around the idea of the nuclear family (Waaldijk 2007, 6). Social rights were conditioned by moral ideals about ideal motherhood, in particular, with mothers being scrutinized, surveilled, and judged based on their "moral behavior" (ibid.). Historically, this led to the normative institutionalization of the "male breadwinner" model (Abramovitz 2018) and the overall racialized model of the nuclear family. Family arrangements that fell outside of this model (e.g., single mother households, families with different cultural backgrounds, or racialized families), revealed the exclusionary mechanisms of state welfare policies. Next to gender, race seems to also play an important role in how these exclusions play out. One example of such a process is the case of African American families in the US context (May 2017). After the civil rights movement, stereotypical and negative representations of the "black matriarch" among others (Collins 2000) were also reproduced by political and policy discourses which represented black women as unfit to be proper mothers and held them responsible for the delinquency of black youth (Moynihan 1965; Toft 2020, 230). Under Ronald Reagan, the trope of the so-called "welfare queen" came to stand for the fraudulent non-white, poor, and young mother that "collect[s] welfare, shunning work and passing on her bad values to her offspring" (Toft 2020, 231).

Similarly to the US context presented earlier, gender, race and family are also intimately connected in the history of Dutch welfare arrangements and its differentiated access. In 2017, the Dutch public was confronted with the *kinderopvangtoeslagaffaire* (childcare benefits scandal), which involved the Dutch tax authorities mistakenly accusing thousands of families of fraudulently claiming childcare allowances (see the example of bias described under the Dutch Tax and Customs Administration in table 12.1). While it provided evidence of explicit targeting of people with a non-Western migrant background (many of them Dutch citizens), historian Chelsea Schields (2023) has noted how the specific association between this category of citizens and the high risk of welfare fraud was not sufficiently discussed. She argues that this link is part of a strong institutional and discursive legacy. Specifically, she traces how welfare retrenchment policies contain normative ideas about family, which emerged

from publicly funded research on Surinamese and Antillean kinship in the 1970s. This research correlates black kinship, single motherhood, and welfare reliance. For Schields, in Dutch welfare institutional practices, "'family becomes racial ontology,' ushering in racist essentialisms through the backdoor of culturalist arguments about family life" (2023, 3). The family becomes metonymic to racial difference, and through this, Dutch public and political institutions locate the origin of social deviance and marginalization in the family and individual behavior, rather than in larger systemic inequalities (p. 18).

Taking cue from Schields' analysis of policy-mediated racial formations through the normativization of the family, we argue that the discriminatory effects of the Rotterdam algorithm for fraud detection, i.e., the targeting of single migrant mothers, are part of a larger racial ontologization institutional process. Emerging discursive practices perpetuate the distinction between the deserving and undeserving subjects of state welfare, which, ultimately manifests in current bureaucratic algorithmization practices. In the next section, we perform a category analysis on the indicators used in the welfare fraud risk-scoring algorithm and place them in this historical context.

Case study: The algorithmic governance of welfare in Rotterdam

In the Netherlands, with its decentralized organization of welfare distribution (Vermeulen 2015), the responsibility for the datafication of welfare distribution and surveillance is delegated to the municipal level. Starting in 2017, the city of Rotterdam decided to implement a data-driven approach to manage the allocation of public benefits. In line with the central government's approach during the same years, which involved experimenting with data analytics techniques to counter welfare fraud, authorities in Rotterdam hired the consulting firm Accenture to develop an automated process that could "identify illegitimate welfare recipients through a truly data-driven approach" (Huyskes 2023). The algorithm built by Accenture was a machine-learning model designed for risk scoring, a popular technique that is often used by banks and financial institutions to assign risk and calculate the trustworthiness of their customers (e.g., their ability to repay a loan or mortgage). In the case of Rotterdam, the purpose was to assess the trustworthiness of welfare recipients and predict fraud risk for each of them.

The Rotterdam welfare fraud risk-scoring algorithm was trained on historical data about known fraud cases, processing 315 variables such as

gender, age, marital and employment status, language skills, neighborhood of residence, number of children, competencies, psychological problems, hobbies, perseverance, the age difference between the children and parents, diplomas and certificates, and several other inputs that represented the subjective evaluations expressed by local caseworkers in each recipient's file (for a full list, see Moti 2021). For instance, one variable pertained to a citizen's availability for appointments with the benefits office and the number of deviated appointments specifically due to their social situation. Another variable describes the ability of a person to deal with pressure and setbacks. By cross-referencing all available data points, each person was assigned a risk score between 0 and 1. Citizens were then sorted by their risk score, resulting in a list of citizens ordered by their "trustworthiness." Those with the highest risk, approximately the top 10 percent, were selected for investigation and surveillance. While this might appear to be a reasonable work practice, the indicator categories used, as well as the measured accuracy of the algorithm, reveal several problems.

First, training an algorithm by finding correlations of known fraud cases with a large number of data categories in order to identify the most relevant indicators is notorious for reproducing past biases (Buolamwini and Gebru 2018). In such cases, it is always unclear how much of the algorithm reflects the entire population of fraudsters and how much the algorithm simply mirrors the personal and institutional focus—in the form of social norms and biases—of past policing practices that created the very specific "sample" on which the algorithm was trained. When we compare all explicitly racialized and gendered indicators in the algorithm with the history of Dutch welfare and its gendered and racializing functions detailed above, we encounter many familiar selection criteria (see table 12.2). By combining these criteria, we can easily distill the type of gendered and racialized identity that is considered most at risk for committing fraud by the municipality of Rotterdam: women with a migration background who became mothers at a young age and have a more than average amount of children. Furthermore, while the respective weights are somewhat on the lower side, the algorithm also takes into account past and current relationships and their length, indicating an interest in family composition and relationship statuses. Again, the added historical context suggests that this is not merely an intrusive government practice resulting from the datafication of welfare, but rather the result of specific past social norms concerning race, gender, and sexuality that were already part of the discourse on welfare and its recipients before its digitization.

Table 12.2. Selection of Indicators from a List of 315 Indicators Used in Richard Moti's Welfare Fraud Risk-Scoring Algorithm

Indicator	Relative importance	Number in list	Explanation	Removed in 2021
Age	100.00	1	Age at time of investigation	No
Language requirement period	15.15	11	Number of days after a person has been assigned a language requirement	Yes
Exemption days for medical reasons	13.82	13	Number of days people have been exempt from applying for jobs for medical reasons	Yes
Length of current relationship	12.44	14	Length of relationship with current partner in days	No
Age at first childbirth	11.17	16	Age of the mother minus the age of her oldest child, which translates to the age of first giving birth.	No
Number of children	10.85	17	Total number of children	No
Sex—woman	9.90	19	Whether or not the person is a woman (0 = No, 1 = Yes)	Yes
Number of young adult children	7.99	26	Total number of young adult children in a household	No
Obstructed due to psychological problems	7.33	28	Hindered from work due to psychological problems	Yes
Language requirement met	4.18	44	Successfully finished a language proficiency course	Yes
Obstructed due to physical problems	3.94	48	Hindered from work due to physical problems	Yes
Spoken language	3.84	49	Spoken Dutch proficiency	Yes
Partner—married	1.10	105	Whether or not a person is married to their partner	No
Number of partners—unmarried	0.45	174	Number of past relationships in the category "unmarried"	No

Source: Compiled from table 2 in the appendix of a letter by Rotterdam Alderman Richard Moti (2021).

Second, while the Rotterdam welfare fraud risk-scoring algorithm is designed to appear as a measurement of trustworthiness, it is, in fact, a crude estimation of risk. This is evident in the fact that almost none of the indicators are directly causally related to fraud but, instead, mostly focus on identity characteristics and contextual factors. Obvious missing indicators are data points related to received forms of welfare (the amounts of money and the

periods over which it was received) and legal entitlement to those welfare payments (whether or not the necessary conditions to receive welfare are met). Apart from a category that mentions whether a person has missed any of their appointments with the Work and Income Bureau, the overwhelming majority of the indicators do not address behavior. To make matters worse, multiple indicators that increase the risk score address whether people have genuine and justified reasons for missing appointments and opportunities for paid labor–such as mental or physical health problems or care responsibilities for children, parents, or other family members. With the choice for identity characteristics and contextual indicators, rather than facts about behavior, the city of Rotterdam has chosen not to police and punish behavior, but aspects of life that people tend to have no choice about.

A final issue with the Rotterdam welfare fraud risk-scoring algorithm is its lack of accuracy—understood as the percentage of correctly predicted fraud cases—which, as noted by Constantaris et al. (2023) based on Rotterdam's internal auditing documents, is "little better than random sampling." Here, we can potentially draw parallels with the aforementioned childcare benefits scandal, which not only exhibited racial bias but, more importantly, produced a significant number of false positives. Both of these cases show that while the assumed potential for accuracy and efficiency is often cited as a reason for initiating the development of an algorithmic system, the actual lack of accuracy and efficiency does not seem to be a deal-breaker once the system is nearing completion or in its implementation phase. While we need more research on the reasons why values such as accuracy and reliability did not seem to be minimum requirements in high-risk government systems, there are indications about the effects of these choices: the unreliability of welfare is deterring people—especially those who often need it the most—from applying for money they rightfully deserve (NOS 2023). This makes contemporary datafied welfare governance a hurdle rather than a last resort for Dutch citizens, whether it was intended to be this, or not.

Conclusion

In this chapter, we discussed how the welfare fraud risk-scoring algorithm used by the city of Rotterdam, roughly between 2018 and 2021, reproduced racist and sexist notions of citizenship. Moreover, we argued that this reproduction should not be seen as an exceptional fact triggered mainly by the current tendencies of datafication and algorithmization of bureaucratic governance. Rather, by explicitly placing contemporary algorithmic practices within their historical

context in the same domain of governance, we showed how racial discrimination and the enforcement of gendered family norms are a continuation of practices that have been a part of the state welfare system for a much longer time.

With this chapter, we aim to contribute to developments in algorithmic auditing and the value-sensitive design of government algorithms. We advocate for an approach that takes into account both the historical contexts and the prevailing discursive practices shaping and reproducing institutional processes. This way, civil servants working on new data projects could potentially know what kinds of inclusion and exclusion they should keep in mind when designing and evaluating their work. In practice, this would mean not only trying to recognize which citizens might be vulnerable now but also which citizens have historically been vulnerable within specific bureaucratic regimes. Quantified auditing techniques, such as the one used by Lighthouse Reports that was detailed in the introduction, are often difficult to perform before an algorithm is put in service. Studying the categories and indicators operationalized in an algorithm and placing them in the historical context of a particular bureaucratic regime as well as national and institutional culture—the type of analysis detailed in this chapter—can be easily done before an algorithm is implemented, potentially saving marginalized citizens from undeserved government scrutiny and governments from making costly mistakes

Acknowledgments

For this work, Gerwin van Schie was supported by a Spinoza grant of the Dutch Research Council (NWO), awarded in 2021 to José van Dijck, professor of media and digital society at Utrecht University. Furthermore, part of the information presented in the case study section was obtained in the context of Diletta Huyskes' research, which is based on qualitative methods, including in-depth interviews with Dutch institutional actors and experts and the analysis of technical documentation obtained by Lighthouse Reports.

References

Abramovitz, Mimi. 2018. *Regulating the Lives of Women: Social Welfare Policy from Colonial Times to the Present.* New York: Routledge.

Anderson, Benedict. 1986. *Imagined Communities: Reflections on the Origin and Spread of Nationalism.* London: Verso.

Aung, Htet, Eva Constantaris, Rik Delhaas, David Davidson, Gabriel Geiger, Ludo Hekman, Daniel Howden, Evaline Schot, and Reinier Tromp. 2021. "Inside a Fraud Prediction Algorithm." Lighthouse Reports. December 18. https://www.lighthousereports.com/investigation/unlocking-a-welfare-fraud-prediction-algorithm/.

Buolamwini, Joy, and Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of Machine Learning Research*, 77–91. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Casimir-Braun, Justin, Eva Constantaras, Htet Aung, Gabriel Geiger, Dhruv Mehrotra, and Daniel Howden. 2023. "Suspicion Machine Methodology." Lighthouse Reports. https://www.lighthousereports.com/suspicion-machines-methodology/.

Choroszewicz, Marta, and Beata Mäihäniemi. 2020. "Developing a Digital Welfare State: Data Protection and the Use of Automated Decision-making in the Public Sector across Six EU Countries." *Global Perspectives* 1(1): 1–15. https://doi.org/10.1525/gp.2020.12910.

Collins, Patricia Hill. 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment.* 2nd ed. New York: Routledge.

Constantaris, Eva, Gabriel Geiger, Justin-Casimir Braun, Dhruv Mehrothra, and Htet Aung. 2023. "Inside the Suspicion Machine." *Wired.* https://www.wired.com/story/welfare-state-algorithms/.

Dencik, Lina. 2022. "The Datafied Welfare State: A Perspective from the UK." In *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, edited by Andreas Hepp, Juliane Jarke, and Leif Kramp, 145–66. Cham: Springer. https://link.springer.com/10.1007/978-3-030-96180-0.

Dencik, Lina, and Anne Kaun. 2020. "Datafication and the Welfare State." *Global Perspectives* 1(1): 1–8. https://doi.org/10.1525/gp.2020.12912.

Entzinger, Han. 2003. "The Rise and Fall of Multiculturalism: The Case of the Netherlands." In *Toward Assimilation and Citizenship: Immigrants in Liberal Nation-States*, edited by Christian Joppke and Ewa Morawska, 59–86. London: Palgrave Macmillan.

Franzke, Aline Shakti. 2022. "An Exploratory Qualitative Analysis of AI Ethics Guidelines." *Journal of Information, Communication and Ethics in Society* 20(4): 401–23. https://doi.org/10.1108/JICES-12-2020-0125.

Franzke, Aline Shakti, Iris Muis, and Mirko Tobias Schäfer. 2021. "Data Ethics Decision Aid (DEDA): A Dialogical Framework for Ethical Inquiry of AI and Data Projects in the Netherlands." *Ethics and Information Technology* 23(3): 551–67. https://doi.org/10.1007/s10676-020-09577-5.

Ghorashi, Halleh, and Ulrike M. Vieten. 2012. "Female Narratives of 'New' Citizens' Belonging(s) and Identities in Europe: Case Studies from the Netherlands and Britain." *Identities* 19(6): 725–741. https://doi.org/10.1080/1070289X.2012.745410.

Gerards, Janneke, Mirko Tobias Schäfer, Arthur Vankan, and Iris Muis. 2022. *Impact Assessment: Fundamental Rights and Algorithms*. The Hague: Ministry of the Interior and Kingdom Relations. https://www.government.nl/documents/reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms.

Heilbron, Belia, and Anouk Kootstra. 2023. "Advocaten: Fraudecontrole DUO Treft Vrijwel Uitsluitend Studenten met Migratieachtergrond." *Investico*, June 21. https://www.platform-investico.nl/artikel/advocaten-fraudecontrole-duo-treft-vrijwel-uitsluitend-studenten-met-migratieachtergrond/.

Huyskes, Diletta. 2023. "'Unprecedented Injustice': Inside the Algorithmization of Social Benefits in the Netherlands." ESPAnet XVI Proceedings: Sistemi di Welfare in Transizione, tra Equità e Sostenibilità. Milan, September 13–15.

Kohler-Hausmann, Julilly. 2007. "'The Crime of Survival': Fraud Prosecutions, Community Surveillance, and the Original 'Welfare Queen.'" *Journal of Social History* 41(2): 329–54.

MacLeavy, Julie. 2016. "Neoliberalism and Welfare." In *Handbook of Neoliberalism*, edited by Simon Springer, Kean Birch, and Julie MacLeavy, 252–61. New York: Routledge.

Maleeyakul, Nalinee, Carola Houtekamer, Merijn Rengers, Gabriel Geiger, Klaas van Dijken, Daniel Howden, Crofton Black, and Ariadne Papagapitos. 2023. "Ethnic Profiling: Whistleblower Reveals Netherlands' Use of Secret and Potentially Illegal Algorithm to Score Visa Applicants." Lighthouse Reports. April 24. https://www.lighthousereports.com/investigation/ethnic-profiling/.

Mann, Monique. 2020. "Technological Politics of Automated Welfare Surveillance: Social (and Data) Justice through Critical Qualitative Inquiry." *Global Perspectives* 1(1): 1–12. https://doi.org/10.1525/gp.2020.12991.

May, Elaine Tyler. 2017. *Homeward Bound: American Families in the Cold War Era.* New York: Basic Books.

Mepschen, Paul. 2012. "Gewone Mensen: Populisme en het Discours van Verdringing in Amsterdam Nieuw-West." *Tijdschrift Sociologie* 8(1): 66–83.

Metaxa, Danaë, Joon Sung Park, Ronald E. Robertson, Karrie Karahalios, Christo Wilson, Jeff Hancock, and Christian Sandvig. 2021. "Auditing Algorithms: Understanding Algorithmic Systems from the Outside In." *Foundations and Trends in Human–Computer Interaction* 14(4): 272–344. https://doi.org/10.1561/1100000083.

Moti, Richard. 2021. "Opvolgen Aanbevelingen Risico-Inschattingsmodel." August 25.

Moynihan, Daniel P. 1965. *The Negro Family: The Case for National Action*. Washington, DC: Office of Policy Planning and Research, US Department of Labor.

NOS. 2023. "Wantrouwen in de Overheid: 'Burgers Zijn Kopschuw Geworden.'" November 15. https://nos.nl/artikel/2497916-wantrouwen-in-de-overheid-burgers-zijn-kopschuw-geworden.

Oudenampsen, Merijn. 2020. "Between Conflict and Consensus: The Dutch Depoliticized Paradigm Shift of the 1980s." *Comparative European Politics* 18: 771–92. https://doi.org/10.1057/s41295-020-00219-0.

Peeters, Rik, and Arjan C. Widlak. 2023. "Administrative Exclusion in the Infrastructure-Level Bureaucracy: The Case of the Dutch Daycare Benefit Scandal." *Public Administration Review* 83(4): 863–77. https://doi.org/10.1111/puar.13615.

Prins, Baukje, and Sawitri Saharso. 2008. "In the Spotlight: A Blessing and a Curse for Immigrant Women in the Netherlands." *Ethnicities* 8(3): 365–84. https://doi.org/10.1177/1468796808092448.

PwC. 2022. *Onderzoek Query's Aan de Poort.* Amsterdam: PwC. https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2022/03/16/onderzoek-querys-aan-de-poort/pwc-onderzoek-querys-aan-de-poort.pdf.

Raji, Inioluwa Deborah, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. "Closing the AI Accountability Gap: Defining an End-to-end Framework for Internal Algorithmic Auditing." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency: FAT* '20*, 33–44. New York: Association for Computing Machinery. https://doi.org/10.1145/3351095.3372873.

Rekenkamer Rotterdam. 2021. *Gekleurde Technologie: Verkenning Ethnisch Gebruik Algoritmes.* Rotterdam: Rekenkamer Rotterdam. https://rekenkamer.rotterdam.nl/wp-content/uploads/2020/11/R.P.20.06-gekleurde-technologie.pdf.

Ruijer, Erna, Gregory Porumbescu, Rebecca Porter, and Suzanne Piotrowski. 2023. "Social Equity in the Data Era: A Systematic Literature Review of Data-Driven Public Service Research." *Public Administration Review* 83(2): 316–32. https://doi.org/10.1111/puar.13585.

Schields, Chelsea. 2023. "A Science of Reform and Retrenchment: Black Kinship Studies, Decolonization, and the Dutch Welfare State." *Contemporary European History* 33(1): 4–22. https://doi.org/10.1017/S0960777323000024.

Siffels, Lotje, David van den Berg, Mirko Tobias Schäfer, and Iris Muis. 2022. "Public Values and Technological Change: Mapping How Municipalities Grapple with Data Ethics." In *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, edited by Andreas Hepp, Juliane Jarke, and Leif Kramp, 243–65. Cham: Springer. https://link.springer.com/10.1007/978-3-030-96180-0.

Tamboukou, Maria. 1999. "Writing Genealogies: An Exploration of Foucault's Strategies for Doing Research." *Discourse* 20(2): 201–17. https://doi.org/10.1080/0159630990200202.

Toft, Jessica. 2020. "History Matters: Racialized Motherhoods and Neoliberalism." *Social Work* 65(3): 225–34. https://doi.org/10.1093/sw/swaa021.

Van Bekkum, Marvin, and Frederik Zuiderveen Borgesius. 2021. "Digital Welfare Fraud Detection and the Dutch SyRI Judgment." *European Journal of Social Security* 23(4): 1–18. https://doi.org/10.1177/13882627211031257.

Van der Woude, Allart, and David Davidson. 2022. "IND Registreerde Afkomst Voor Risicoprofilering." *VPRO*. https://www.vpro.nl/argos/lees/onderwerpen/artikelen/2022/ind-registreerde-afkomst-voor-risicoprofilering.html.

Van Gent, Wouter, Cody Hochstenbach, and Justus Uitermark. 2018. "Exclusion as Urban Policy: The Dutch 'Act on Extraordinary Measures for Urban Problems.'" *Urban Studies* 55(11): 2337–53. https://doi.org/10.1177/0042098017717214.

Van Gerven, Minna. 2019. "The Dutch Participatory State: Shift from a Welfare System of Collective Solidarity towards Individual Responsibility in a Participatory Society." In *Routledge Handbook of European Welfare Systems*, edited by Sonja Blum, Johanna Kuhlmann, and Klaus Schubert. Milton: Routledge.

Van Schie, Gerwin, Alex Smit, and Nicolás López Coombs. 2020. "Racing through the Dutch Governmental Data Assemblage: A Postcolonial Data Studies Approach." *Global Perspectives* 1(1): 1–19. https://doi.org/10.1525/gp.2020.12779.

Van Schie, Gerwin, and Serena Oosterloo. 2020. "Predictive Policing in the Netherlands: A Critical Data Studies Approach." In *A Critical Approach to Police Science: New Perspectives in Post-Transitional Policing Studies*, edited by Kerezsi Klára and Veronika Nagy, 169–96. The Hague: Eleven International.

Van Zoonen, Liesbet. 2020. "Data Governance and Citizen Participation in the Digital Welfare State." *Data & Policy* 2(e10): 1–17. https://doi.org/10.1017/dap.2020.10.

Vecchione, Briana, Karen Levy, and Solon Barocas. 2021. "Algorithmic Auditing and Social Justice: Lessons from the History of Audit Studies." In *Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–9. New York: Association for Computing Machinery. https://doi.org/10.1145/3465416.3483294.

Vermeulen, Wouter. 2015. "Decentralization of Social Policy in the Netherlands." The Hague: CPB. https://www.cpb.nl/sites/default/files/publicaties/download/cpb-background-document-decentralization-social-policy-netherlands.pdf.

Waaldijk, Berteke. 2007. "Beyond Social Citizenship: New Approaches in Comparative European Welfare History." In *Reciprocity and Redistribution: Work and Welfare Reconsidered*, edited by Gro Hagemann. Edizioni Plus.

Wieringa, Maranke. 2020. "What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 17. Barcelona: ACM. https://doi.org/10.1145/3351095.3372833.

Yanow, Dvora, Marleen van der Haar, and Karlijn Völke. 2016. "Troubled Taxonomies and the Calculating State: Everyday Categorizing and 'Race-Ethnicity'—The Netherlands Case." *The Journal of Race, Ethnicity, and Politics* 1(2): 187–226. https://doi.org/10.1017/rep.2016.7.

About the Authors

Gerwin van Schie is assistant professor of new media and digital culture at Utrecht University.

Laura Candidatu is assistant professor in gender studies at Utrecht University.

Diletta Huyskes is a PhD candidate in sociology at the University of Milan and Turin.