



**The Datafication of Race-Ethnicity**

Gerwin van Schie

# **The Datafication of Race-Ethnicity**

An Investigation into Technologically Mediated Racialization  
in Dutch Governmental Data Systems and Infrastructures

Gerwin van Schie

The cover was designed by Christl de Kloe. It features a part of the “Portrait of the Global Citizen”, a visualization of parts of human DNA hanging on the walls of the Green Salon in the Dutch royal palace “Huis ten Bosch”. The photo used was made by Sylvain Deleu and is available for reuse under a creative commons license (CC0) on the website of the Central Government Real Estate Agency (see Rijksvastgoedbedrijf 2019).

© 2022 Gerwin van Schie

All rights reserved

ISBN: 978-94-6361-734-5

# The Datafication of Race-Ethnicity

An Investigation into Technologically Mediated Racialization  
in Dutch Governmental Data Systems and Infrastructures

## De dataficatie van ras-ethniciteit

Een onderzoek naar technologisch gemedieerde racialisatie in  
datasystemen en -infrastructuren van de Nederlandse Overheid  
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de  
Universiteit Utrecht  
op gezag van de  
rector magnificus, prof.dr. H.R.B.M. Kummeling,  
ingevolge het besluit van het college voor promoties  
in het openbaar te verdedigen op

vrijdag 16 september 2022 des middags te 4.15 uur

door

Gerardus Adrianus van Schie

geboren op 2 januari 1987  
te Leiden

Promotoren:

Prof. dr. S. Ponzanesi

Prof. dr. J.F.T.M. van Dijck

Beoordelingscommissie:

Prof. dr. W.H.K. Chun

Prof. dr. E. Isin

Prof. dr. A.A. M'Charek

Prof. dr. H.O. Dijkstra

Prof. dr. A.J. Meijer (voorzitter)

Dit proefschrift werd mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

# Contents

<b>List of Figures</b> .....	<b>VII</b>
<b>Acknowledgements: “A Good Infrastructure Is Hard to Find”</b> .....	<b>IX</b>
<b>Preface</b> .....	<b>XIII</b>
<b>“Pulled from the Clay”: An Introduction to the Datafication of Race-Ethnicity in the Netherlands</b> .....	<b>1</b>
Datafication .....	2
Race-Ethnicity as “Blood and Soil” .....	3
Research Questions .....	9
On Positionality.....	11
Chapters and Case Studies.....	14
<b>Chapter 1: Understanding Technologically Mediated Racialization</b> .....	<b>19</b>
<b>Racialization</b> .....	<b>20</b>
<b>Technologically Mediated Racialization</b> .....	<b>28</b>
<i>The Instrumental Approach – Racialization as Content</i> .....	28
<i>The Epistemological Approach – Racialization as Operation</i> .....	32
<i>The Ontological Approach – Towards a Technicity of Race</i> .....	41
<b>Conclusion</b> .....	<b>45</b>
<b>Chapter 2: Investigating Technologically Mediated Racialization in Data Systems</b> .....	<b>47</b>
<b>Fields of Study</b> .....	<b>47</b>
<i>Critical Data Studies</i> .....	47
<i>Postcolonial Studies</i> .....	49
<i>Interventions</i> .....	51
<b>Methods</b> .....	<b>53</b>
<i>Critical Discourse Analysis</i> .....	54
<i>Affordance Analysis</i> .....	56
<i>Infrastructural Inversion</i> .....	58
<b>Conclusion</b> .....	<b>60</b>
<b>Chapter 3: The Origins of Race-Ethnic Categorization in the Dutch Governmental Data Ontology (1899-2020)</b> .....	<b>65</b>
<b>Social Stratification through Pillarization (1899–1940)</b> .....	<b>70</b>
<i>“Nationality” and “Place of Birth”</i> .....	71
<i>Pillarization and the Racialization of Religious Affiliation</i> .....	74
<i>Implications</i> .....	77
<b>The Invention of “Autochtoon” and “Allochtoon” (1945–1971)</b> .....	<b>78</b>
<i>“Nationality” and “Place of Birth”</i> .....	81
<i>Moving from Pillarization towards Multiculturalism</i> .....	85

<b>The Datafication of Race-ethnicity (1980–2021)</b> .....	<b>88</b>
<i>Dutch “Foreigners”</i> .....	91
<i>Dutch Technologically Mediated Racialization</i> .....	93
<i>“Migration Background” and Colonial Aphasia</i> .....	95
<b>Conclusion</b> .....	<b>99</b>
<b>Chapter 4: The Appropriation of Race-Ethnic Data in Governmental and Non-Governmental Data Applications</b> .....	<b>103</b>
<b>Case-study 1: The Allochtonenmeter</b> .....	<b>107</b>
<i>Technologically Mediated Racialization in the Allochtonenmeter</i> .....	114
<b>Case-study 2: The Leefbaarometer</b> .....	<b>116</b>
<i>Visualizing Livability</i> .....	118
<i>Population Indicators</i> .....	124
<i>The Leefbaarometer and the Rotterdam Act as Technological Redlining</i> .....	128
<i>Technologically Mediated Racialization in the Leefbaarometer</i> .....	130
<b>Conclusion</b> .....	<b>131</b>
<b>Chapter 5: Race-ethnicity and the Datafication of Policing</b> .....	<b>135</b>
<b>Case Study 3: The “Crime Anticipation System” of the Dutch national police</b> .....	<b>138</b>
<i>Data and Infrastructures</i> .....	140
<i>Demographic Indicators</i> .....	144
<i>From “Known Offenders” to “Suspects”</i> .....	146
<i>Workflow</i> .....	147
<b>CAS in the Context of Ethnic Profiling in the Netherlands</b> .....	<b>150</b>
<b>Predictive Policing and Technologically Mediated Racialization</b> .....	<b>152</b>
<b>Conclusion</b> .....	<b>156</b>
<b>Chapter 6: Conclusions and Recommendations</b> .....	<b>159</b>
<b>Findings</b> .....	<b>159</b>
<i>Categorization</i> .....	160
<i>Datafication</i> .....	161
<i>Appropriation</i> .....	162
<i>The Practices and Consequences of the Datafication of Race-ethnicity</i> .....	164
<b>Governing Race-Ethnic Categorizations in a Datafied Society</b> .....	<b>168</b>
<i>Classification and Race-ethnic Determination/Self-Determination</i> .....	168
<i>Clustering, Labeling, and Situatedness</i> .....	169
<i>Datafication and Accountability - Reversing the Order</i> .....	170
<b>Further Research on Technologically Mediated Racialization</b> .....	<b>174</b>
<b>Bibliography</b> .....	<b>177</b>
<b>Summary in English</b> .....	<b>201</b>
<b>Samenvatting in het Nederlands</b> .....	<b>203</b>
<b>Curriculum Vitae</b> .....	<b>206</b>

# List of Figures

FIGURE 1: PARTIAL INTERIOR OF THE "GROENE SALON". PHOTO BY CORNÉ BASTIAANSEN.....	1
FIGURE 2: CELEBRATION OF SINT NIKOLAAS AT THE VAN SCHIE FAMILY, DECEMBER 1988. ....	12
FIGURE 3: THREE APPROACHES TO TECHNOLOGICALLY MEDIATED RACIALIZATION .....	46
FIGURE 4: FUNCTION OF THE THREE MAIN METHODS IN EACH CASE STUDY.....	63
FIGURE 5: "NATIONALITY", COMPILED FROM THE CENSUS REPORTS OF 1899, 1909, 1920, AND 1930 .....	72
FIGURE 6: "PLACE OF BIRTH", COMPILED FROM THE CENSUS REPORTS OF 1899, 1909, 1920, AND 1930. ....	73
FIGURE 7: EXCERPT OF CENSUS REPORT OF 1909, FEATURING PART OF PAGE 242. ....	74
FIGURE 8: EXCERPT OF CENSUS REPORT OF 1909, FEATURING PART OF PAGE 243 .....	75
FIGURE 9: LEFT SIDE OF TABLE WITH DIFFERENT RELIGIOUS DENOMINATIONS IN THE NETHERLANDS .....	76
FIGURE 10: RIGHT SIDE OF DIFFERENT RELIGIOUS DENOMINATIONS IN THE NETHERLANDS .....	77
FIGURE 11: PROTEST AGAINST THE 1970 CENSUS .....	80
FIGURE 12: PLACE OF BIRTH IN THE CENSUSES OF 1947, 1960, AND 1970.....	81
FIGURE 13: NATIONALITIES AS FEATURED IN THE 1947 CENSUS. ....	83
FIGURE 14: NATIONALITIES AS FEATURED IN THE 1971 CENSUS. ....	84
FIGURE 15: SCHEMATIC OVERVIEW OF DECISION TREE FOR DETERMINING A DUTCH CITIZEN'S 'ETHNIC' ORIGINS...	90
FIGURE 16: CLUSTERING OF DUTCH CITIZENS BASED ON COUNTRY OF ORIGIN. ....	91
FIGURE 17: PROPORTION OF THE "FOREIGN POPULATION" IN THE POTENTIAL WORKFORCE IN THE 2001 CENSUS. .	92
FIGURE 18: "DURATION OF STAY IN THE NETHERLANDS" IN THE 2011 CENSUS. ....	93
FIGURE 19: CBS TABLE SORTED BY MIGRATION BACKGROUND.....	96
FIGURE 20: FINDINGS IN DIFFERENT TIME PERIODS DIVIDED BY APPROACH .....	101
FIGURE 21: THE INTERFACE OF THE ALLOCHTOON-O-METER AS SHOWN ON THE GEENSTIJL WEBSITE IN 2017. ....	108
FIGURE 22: SCREENSHOT OF THE ALLOCHTONENMETER MADE IN 2012 .....	110
FIGURE 23: SCREENSHOT OF THE ALLOCHTONENMETER MADE IN 2018 .....	111
FIGURE 24: SCREENSHOT OF THE ALLOCHTONENMETER MADE IN 2021 .....	112
FIGURE 25: EQUIVALENTS OF RACE-ETHNIC CATEGORIZATION IN OLD AND NEW VOCABULARY.....	113
FIGURE 26: MAP OF UTRECHT, THE NETHERLANDS, IN THE LEEFBAAROMETER APP, AT LOW RESOLUTION. ....	117
FIGURE 27: MAP OF UTRECHT, THE NETHERLANDS, IN THE LEEFBAAROMETER APP, AT MEDIUM RESOLUTION .....	122
FIGURE 28: MAP OF UTRECHT, THE NETHERLANDS, IN THE LEEFBAAROMETER APP, AT HIGH RESOLUTION .....	122
FIGURE 29: THE LEEFBAAROMETER INFORMATION MENU .....	123
FIGURE 30: DIMENSIONS WITH PERCENTAGES IN THE ALGORITHM USED IN THE CALCULATION OF LIVABILITY .....	124
FIGURE 31: INDICATOR TABLE OF THE "INHABITANTS DIMENSION" OF THE LEEFBAAROMETER .....	125
FIGURE 32: CYCLE OF INFORMATION IN THE CAS DATA ASSEMBLAGE.....	141
FIGURE 33: INTERCONNECTIONS BETWEEN DIFFERENT DATABASES USED FOR CAS .....	141
FIGURE 34: INDICATORS USED IN CAS, ACQUIRED FROM CBS. ....	145
FIGURE 35: BVI INDICATORS USED FOR EACH CAS SQUARE. ....	147
FIGURE 36: PICTURE OF THE INTERFACE OF CAS.....	148
FIGURE 37: PART OF THE DECISION-MAKING PROCESS OF DUTCH POLICE OFFICERS.....	151





## Acknowledgements: “A Good Infrastructure Is Hard to Find”

While this dissertation will undoubtedly herald a new phase in my life, it is also the capstone to a few years of hard work. To get to this point, I had to navigate many different institutional and physical infrastructures while meeting many different people along the way. The quote “a good infrastructure is hard to find” is from Geoffrey Bowker and Susan Leigh Starr’s *Sorting Things Out* (1999, 33), a study that was instrumental to my research. I find its dual meaning very interesting: the moral, emotional, financial, and technical support that I was lucky enough to receive is quite scarce and such support tends to “fade into the woodworks” (Bowker and Star 1999, 34) once a project, such as this one, has been completed. In these acknowledgements, I would, therefore, like make visible the various supportive foundations that enabled me to write this monograph.

While it might sound surprising in hindsight, the beginning of my academic journey was not entirely successful. I failed four times at making it past the first year of a bachelor program. After having tried my luck at Artificial Intelligence, Philosophy, Business Administration, and Philosophy and Religion Education respectively, I ended up in the BA program of Theater, Film, and Television Sciences at Utrecht University in 2011. This time I did much better: I actually managed to graduate. However, there is more to this story than mere personal perseverance; I could not have succeeded without the privilege of receiving financial support both from family and the Dutch government. Not everybody is given a second chance at a university education, let alone a third, fourth, and fifth.

Given my interest in new media, I decided to do an MA in New Media and Digital Culture, also taught at the Media and Culture Department of Utrecht University. In this program, I met many inspiring teachers who would later become colleagues, such as René Glas, Imar de Vries, Stefan Werning, Michiel de Lange, and Joost Raessens. I would like to give a special mention to Marianne van den Boomen, who came up with an idea for my research internship only weeks before she passed away.

I am thankful for the opportunity to do this research internship with Mirko Schäfer at the Utrecht Data School (UDS) in 2014. I still have fond memories of working together with Irene Westra in the small office in the attic of Kromme Nieuwegracht 20. We decided to take a few weeks to figure out both how to scrape an online forum *and* figure out the ethical implications of such a method. Inexperienced as we were, we followed the wrong order and found out that the technical part of our research was not as difficult as we had envisaged. I will never forget how, after three days into this investigation, we had to tell Mirko we managed to download information concerning more than 130,000 people without having read up on research ethics yet. Mirko, thank you so much for turning these bad research practices into an opportunity to write about big data research ethics and for preparing us for our first conference presentation and our first publication. I am also very

grateful that you, while consistently advising me against an academic career, still provided me with many opportunities to gain experience as a young pre-PhD researcher at UDS while actually being paid, an uncommon feat in the Humanities. Moreover, I would like to thank Aline Franzke, Sander Prins, Daniela van Geenen, James Dyer, Iris Muis, Marjolein Krijgsman, Joris Veerbeek, Lisa de Graaf, Max Boiten, and all the other people I had the pleasure of working and meeting with at UDS before and during my PhD. Those lunchbreaks, drinks, discussions, and Among Us sessions were always a source of joy and inspiration.

While finishing my MA degree and working for UDS, a romantic relationship led me to spend many months in Turkey. While there, I could look at the Netherlands from a distance and learn from someone dear to me what it is like to be Turkish in the Netherlands, and Kurdish in Turkey. Thank you Zeynep, conversations about these topics were very important for my growing understanding of the arbitrary nature of race-ethnic categories and their discriminatory effects. These insights, together with the knowledge I gained at UDS, were instrumental in writing the PhD-proposal that led to this dissertation.

I would like to thank the Dutch Research Council (NWO) for providing me with funding for four years of PhD research through their “PhDs in the Humanities” talent grant scheme. Furthermore, I am very grateful to Frans Ruiters, Frank Kessler, and Hanneke Jansen from ICON, who provided feedback and coaching throughout the stressful application procedure and during my time as a PhD. While researching, I was happy to be part of the NOG research school. Thank you, Rosemarie Buikema, Berteke Waaldijk, and Trude Oorschot for organizing the reading and writing seminars, mock defenses, and for taking the time to comment on the work of young scholars who are still perfecting their craft. Furthermore, Iris van der Tuin, Karin van Es, Ingrid Hoofd, Koen Leurs, Veronika Nagy, thank you for giving feedback on early versions of chapters, and for giving me the opportunity to present my work and sharpen my argumentation in guest lectures in your courses.

I would like to thank Stephanie de Smale, Desi Prianti, Laura Candidatu, Adriano José Habed, Gianmaria Colpani, Katrine Smiet, Zerrin Cengiz, Donya Alinejad, Nico de Klerk, Dulce da Rocha Gonçalves, Jamilla Notebaard, Javier Zamora García, Duan Shichang, Irene Alcubilla Troughton, Dennis Jansen, Wouter Egelmeers and all the other people who frequented Drift 15 for their support and our many pleasant conversations about theory, work, labour conditions in academia, and life in general. Similarly, I want to thank Viktorija Morozovaite, Anna Marieke Weerdmeester, Fernando van der Vlist, Laura Frederika Lalíková, and all the other PhD’s from Governing the Digital Society for the fun we had during lunch at José van Dijck’s office and drinks at the Bastaard. I would like to thank Maranke Wieringa, Rosa Wevers, and Tim de Winkel, in particular, for our collaborations including frequent exchanges of writing samples, feedback, and discussions. Furthermore, thank you Serena Oosterloo, Nicolás López Coombs, Alex Smit, Marco van der Vinne,

Laura Dragulin, and Araa al Jaramani for collaborating with me in research on one of the case studies in this dissertation, and other related projects.

I want to thank Sandra Ponzanesi and José van Dijck for their outstanding supervision. Looking back, I find it remarkable how you were able to both temper my optimism and overconfidence at the beginning of my PhD, and how you supported and encouraged me when my final two years happened to coincide with a worldwide pandemic. Thank you, Christien Franken, not only for your editing work, but also for the much-needed coaching sessions when I was struggling to get my thoughts on paper in a coherent manner.

It is important to have friends you can rely on especially during an at times mentally taxing project such as this one. Thank you Yvette, Robin, Tess, Yentl, Tamara, Djarno, Tim, Lotje, and James for the shared holidays and other good times we had together over the past decade. You are all amazing people and I genuinely hope that finishing this PhD will mean that I have much more time to spend with you on trips and nights out!

Finally, I want to thank my family: my father and sisters for their never-ending support, despite my rocky start in academia, and the home front, consisting of Cixous the cat and my life partner, the most inspiring person I have ever met. Christl, thank you for always having my back, not only during the mundane and everyday tasks, but also in frequently discussing, and critiquing the theoretical and analytical parts of this dissertation while I was writing it. I could not have done this without you! I'm looking forward to the many adventures we will have in the future, and sincerely hope that I, some day, may return the favour.



## Preface

The research presented in this dissertation is the result of the NWO funded project “Datafication of Race and Ethnicity in the Netherlands: Investigating Practices, Politics and Appropriation of Governmental Open Data” that was carried out between 2017 and 2022.

Several early versions and parts of the different case studies and chapters in this dissertation have been published previously as journal articles and book chapters. An early version of Chapter 3 was published as “Origins: A History of Race-Ethnic Categorisation in the Dutch Governmental Data Ontology (1899–2018)” in *TMG - Journal for Media History* (see van Schie 2018). Parts of the journal article “Racing through the Dutch Governmental Data Assemblage: A Postcolonial Data Studies Approach” published in *Global Perspectives* (see van Schie, Smit, and López Coombs 2020) can be found in Chapter 1 and in the case studies of Chapter 4. Furthermore, early versions of Chapter 5 were published as the article “The Politics and Biases of the ‘Crime Anticipation System’ of the Dutch Police” in the *Proceedings of the Workshop on Bias in Information, Algorithms and Systems* (see Oosterloo and van Schie 2018), and the book chapter “Predictive Policing in The Netherlands: A Critical Data Studies Approach” in *A Critical Approach to Police Science: New Perspectives in Post-Transitional Policing Studies* (see van Schie and Oosterloo 2020). Finally, parts of the book chapter “The Datafication of Racialization and the Pursuit of Equality: The Case of the “Barometer Culturele Diversiteit” from the edited volume *Situating Data: A Cultural Inquiry* (see van Schie forthcoming) can be found in Chapter 1 and Chapter 6.



## “Pulled from the Clay”: An Introduction to the Datafication of Race-Ethnicity in the Netherlands



Figure 1: Partial interior of the "Groene Salon". Photo by Corné Bastiaansen.

In 2019, *Huis ten Bosch*, home to the Dutch King Willem Alexander and Queen Maxima van Oranje, was reopened after a four-year renovation project. Jacob van Breugel, a world-renowned artist with “Dutch roots”, made five wall-to-wall artworks for the so-called “*Groene Salon*” of Huis ten Bosch (see Kleijn 2019). Intended as modern takes on the age-old painted portrait, the installed artworks are large abstract datafied representations of particular details of the DNA of the King and Queen, which were retrieved and sequenced with the help of the molecular geneticist and former president of the Royal Netherlands Academy of the Arts and Sciences, Hans Clevers. The two artworks depicting the King and Queen are placed in a traditional manner, featuring their DNA-portraits on opposing sides of the room. The remaining spaces on the walls are filled with three other portraits titled “Portrait of the Past”, “Portrait of the Global Citizen”, and “Portrait of the Future”. Each of these five artworks features a similar visualization made up of a collection of neat rows of white and red bricks, which are partly made from Rhine river clay. The coloured bricks alternate in a seemingly random fashion, reminiscent of a collection of bar codes (see Cover image and Figure 1). Because the artworks are on all three walls around the



room (the side with the large windows excluded), the visitor is meant to experience a feeling of “immersion”, not unlike the feelings invoked in a so-called *Gesamtkunstwerk* (Kleijn 2019, 26). The mixture of incoming light and the colours of the bricks cause the pattern to give off an orange golden glow, a direct reference to the family name of the Dutch royals: Van Oranje.

While the wall covering of *De Groene Salon* looks like a novel and intriguing interplay between traditional elements of typical art featured in a royal palace and state-of-the-art DNA-technology, it combines two ideas that I problematize throughout this dissertation: 1) the idea that new technologies that quantify people in the form of data are objective knowledge-producing instruments, and 2) the idea that nationality, in this case Dutchness, is a biological and geographical concept, i.e. tied to “blood” and “soil”. In this introduction, I will use the artwork as an abstract graphic example to introduce the main concepts and questions that guide my investigation of what I will call “technologically mediated racialization”. Furthermore, as the location of the artwork suggests, the views they represent are not opinions of single people in the margins. Rather, they are representative of institutionally embedded ideas of the Dutch approach to data technologies and to race.

## Datafication

The first impression of the DNA-portraits, which look like a collection of barcodes only readable by a computer, is highly reflective of our current age of datafication. Datafication is understood in this dissertation as the transformation of all aspects of life into quantifiable data (Mayer-Schönberger and Cukier 2013). While processes of datafication are manifest in most areas of business and science, the quantification of people’s bodies and their behaviour has been at the core of this development from the onset (Gitelman 2013; Browne 2015; Couldry and Mejias 2019b). The perception of datafied and computational technologies as the apex of scientificity and objectivity (see boyd and Crawford 2012; Kitchin 2014b) is, however, “staked in ideological assumptions, which are, in turn, rooted in prevailing social norms” (Van Dijck 2014, 200). Despite their appearance, data are, therefore, never raw, but always already “cooked” (Bowker 2013). Data make particular things visible, but also actively silence other ways of knowing, and rarely show what has been excluded. It has been suggested that the word “data” itself, with its meaning of “given”,<sup>1</sup> is obfuscating that data are in actuality “taken” and should, therefore, be called “capta” (see Drucker 2011; Kitchin 2014b, 2). As a result, we should always understand that data are produced

---

<sup>1</sup> This etymology can be clearly seen in the Dutch language in the word *gegeven*, which can literally be translated as “given”. In the Netherlands, the plural *gegevens*, is still commonly used to designate digital data.

by a multiplicity of entangled forces and never simply convey the purported referent. There is always already an excess, a beyond, a mutually constituted exclusion, and an unconscious history of data. (Dixon-Román 2017, 45)

The aim of this dissertation is, therefore, to pay attention to what is taken, but also to what is excluded, and retrieve the “excess” and “unconscious history of data” by tracing its connections and reconstructing the lost connections that data are haunted by.

Like the data systems under scrutiny in this dissertation, the DNA-portraits mentioned above are haunted by discourses concerning the national, ethnic, and racial heritage of the Dutch. To have a chance at eliminating the problematic and negative effects of these discourses from contemporary datafied instruments, “we must learn how to identify hauntings and reckon with ghosts, must learn how to make contact with what is without doubt often painful, difficult, and unsettling” (Gordon 1997, 23). In the words of feminist literature and art scholar Rosemarie Buikema:

Each text and every sign we use for signifying what we mean and intend has been used before, has figured in another context. On the one hand, it is the prerequisite for making ourselves understood; on the other hand, the connotations of that former usage are brought along in the linguistic sign. Because of all those intended and unintended connotations, we usually say more, by means of words and images, than we are aware of. (Buikema 2009, 73)

Such “intended and unintended connotations”, especially in the case of topics such as ethnicity, race, and nationality, are not only communicated by images and text, but can just as well be communicated through the material and informational qualities of an artwork (see Buikema 2009; 2012; Wevers 2018). By paying close attention to the information that is communicated and the way in which it is visualized, we can retrace the connections, unconscious histories, and hauntings that connect contemporary practices of datafication in the Netherlands with the country’s past discourses on religious segregation, discrimination, and colonialism.

### **Race-Ethnicity as “Blood and Soil”**

One of the unconscious histories of DNA, when understood as data pertaining to the ancestry of a person, is its haunting by both cultural and biological conceptualizations of race (see M’charek 2000; 2005; Saini 2020). This long history can be recognized in the connection of the materials used for the graphic visualization of the DNA-portraits with age-old discourses on race, ethnicity, and nationality. For example, the genetic information obtained from the King and Queen is quite literally, in the words of Jacob van Breugel, “*uit de klei getrokken*” or “pulled from the clay” (Kleijn 2019, 26). While this Dutch saying can be used to describe a person or their behaviour as *boers*,

or “unaccultured”<sup>2</sup>, in this artwork it refers to the clay from which the bricks were made by the artist. Van Breugel mixed industrial clay and Rhine river clay to turn its original red colour, which he describes as “like blood: the most primitive element we all have in common” (Kleijn 2019, 27), into a softer orange to signify the last name of the royal family: *van Oranje*. In this way, the combination of the material, its colour, and the information represented in the DNA-portraits of the King and Queen quite neatly tie together several related, long-standing ideas about the connection of blood, soil, and nation.

The use of Dutch soil in the form of clay in the representation of the genetic code of the royal family resonates with the Dutch vocabulary that was used to categorize people as Dutch or non-Dutch in terms of their ancestry for approximately 50 years. In the 1970s, the Dutch government began to use the words *allochtoon* and *autochtoon*, made up from the Greek word *chtōn*, meaning “earth” or “soil”, and the prefixes *auto* and *allo*, meaning the “same” and “different” respectively, to distinguish between people that were considered to have Dutch ancestry and people that were considered to have a migration background (Yanow and Van der Haar 2013, 237). The distinction used to assign these categories was made on the basis of the birthplace of the biological parents of a person: with two parents born in the Netherlands, someone was considered to be “autochtoon”; with one parent born in another country, someone was considered to be “allochtoon”. In this regard, it combined two common ways of assigning someone a nationality: either based on the nation, on the soil of which they are born, called *ius soli*, or based on the nationality of their bloodline, called *ius sanguinis*. In the case of the Netherlands, the “allochtoon”/“autochtoon” dichotomy should, however, be seen not as a distinction addressing legal formal citizenship, but rather as way to ascribe a form of moral citizenship, “an *extra-legal normative concept*” (Schinkel 2010, 167, italics in original) of who count as proper citizens of The Netherlands and who do not. While the citizenship discourses on blood and soil seem to express opposing ideas, they are actually informed by a similar principle based on birthright, silently implying that a persons choice for a new home country, for whatever reason, is merely a cosmetic one:

Both parentage and territorial principles rely upon circumstances of birth as the main criteria for distinguishing insiders from outsiders. In other words, both membership-transfer principles are ascriptive in nature; they remain caught in the territory/blood trap, despite their very different narratives of nationhood and statehood. (Shachar 2009, 127)

By using Dutch soil to tie the bloodline of the Dutch royal family to the Dutch nation, the artwork in *De Groene Salon* is reifying the idea that membership of The Netherlands is first and foremost

---

<sup>2</sup> The literal translation of *boers* would be “farmerlike”.

based on blood and soil, rather than for instance on free choice or political and ideological affiliation.

If we go even further down the trail of blood and soil that is invoked both by the artwork of Van Breugel and governmental racial categorizations, we encounter a more sinister past. While the discursive connection of blood and soil has existed in Europe from prehistoric ages (see Linke 1999), in Dutch cultural memory, it is most prominently tied to the German words “*blut und boden*,” signifying the cultural and biological policies of Nazi Germany in the 1930s and 1940s (see Lovin 1967). Of course, the most well-known book that explained these ideas, namely *Neuadel aus Blut und Boden* (1930) from the Nazi ideologue Richard Walther Darré, can hardly be considered Dutch. However, the far right of the political spectrum in the Netherlands before and during WW II developed its ideology in close relation with German discourse (see Orlow 2016). Therefore, the Dutch words “*bloed en bodem*” too are connected in particular to Dutch Nazi discourses on topics such as politics (Kunkeler 2018, 209–10), art (Mulder 1982; Ensel 2013), eugenics (Snelders 2007), and colonialism (von Frijtag Drabbe Künzel 2017). One prime example of the Dutch discourse on race during WW II, is the book *Wien Neerlandsch Bloed...: Het Rassenvraagstuk en zijn Beteekenis voor Nederland* (van Schoping 1941)<sup>3</sup> which discusses the practical problems of classifying people based on their biological ancestry as follows:

The most radical point of view—and essentially the old Nordic standpoint—would be to recognize as Dutch, only those who are of pure blood, and thus draw the line where foreign blood is present, even if its percentage is very small. (van Schoping 1941, 119; my translation)<sup>4</sup>

This statement resonates with the infamous principle of hypodescent, colloquially referred to as the “one drop” rule, that was used in the United States until late in the 20<sup>th</sup> century, which meant that if a person had only one black ancestor they were considered black, no matter how far back in the family tree the ancestor could be found (see Omi 1997, 8; Denton 1997, 87). While such policies might seem preposterous from a contemporary Dutch perspective, many of the underlying principles of this classification practice survive until today in Dutch governmental data

---

<sup>3</sup> This title can be translated as *Of Dutch Blood... The Race Question and its Significance for The Netherlands*. The first part of the title is a reference to the former Dutch anthem called “Wien Neerlands Bloed”. The name of the writer, Van Schoping, was a pseudonym used by F.E. Farwerck who was a close friend of Anton Mussert, the founder of the Dutch National Socialist movement (see H. Hoogenboom 2015).

<sup>4</sup> The original Dutch text: “Het meest radicale standpunt—en in wezen het oude Noordrasstandpunt—zou zijn, alleen degenen, die van zuiveren bloede zijn, als Nederlanders te erkennen en de streep dus daar te trekken, waar vreemd bloed, als is het percentage nog zoo klein, aanwezig is” (van Schoping 1941, 119).

systems. Statistics Netherlands<sup>5</sup> (*Centraal Bureau voor de Statistiek*, in short: CBS) still classifies people based on the birthplace of their parents, not their own. Furthermore, the category “autochtoon” that was used for people with Dutch parents was effectively kept pure by not permitting people with foreign parental blood to be classified as such, in some cases until the third generation after migration (see, for example, CBS 2016). By explicitly connecting ideas about ancestry, blood, and soil, the DNA-portraits seem to be haunted by ideas that not only inform classification practices of the Dutch government, but also ideas that are connected with all kinds of discourses on race and colonialism.

As will hopefully be clear from the above, the portraits of this King and Queen offer interesting and problematic discussion material, not in the least because of their biologically essentialist representations of masculinity and femininity. For the next part, however, I will focus on another portrait on the wall of *De Groene Salon*, namely “The Portrait of the Global Citizen”. Intended as a portrait of “mankind”, this piece is made up of DNA that is present in “all of us” (Van Breugel in Kleijn 2019, 28). The data used for this portrait consists of several different pieces of DNA, namely, HACNS1 (responsible for the development of the opposable thumb), FOXP2 (responsible for the development of speech), ARHGAP11B (responsible for brain growth in comparison to other primates), and CREB1 and CREB2 (responsible for the development of long-term memory) (Kleijn 2019, 30). While these first five parts of DNA are easily recognizable as universal human characteristics, the characteristic of the sixth, LCT gene, is that it allows “humans – particularly in Northwest Europe –[to] continue to drink milk and digest lactose past infancy” (Kleijn 2019, 28). While the “nonmutated LCT gene” is also represented in the artwork it is still peculiar that this piece of DNA is considered significant in a portrait that is supposed to refer to the “global citizen”. In an interview, geneticist Hans Clevers explains how the mutated gene is particularly significant for the Dutch, as it symbolizes their ability to “drink milk and eat cheese”, while the non-mutated gene, which is related to lactose-intolerance in adulthood, is referred to as “African” (NOS 2019). Such a use of milk as a symbol to distinguish between “Western” and “non-Western” people resonates with the way in which it has become a recurring theme in white supremacist circles. References to dairy milk have become an important part of the construction of white masculinities and their opposition to the colonial Other in online alt-right discourses (Gambert and Linné 2018). Additionally, for white supremacists, the practice of drinking copious amounts of milk during protests and counterprotests has become a way of demonstrating their superiority

---

<sup>5</sup> In The Netherlands, Statistics Netherlands is called *Centraal Bureau voor de Statistiek*. The literal translation of this name would be “Central Bureau for Statistics.” Since this rather well-known institution is usually referred to by means of the abbreviation CBS, for reasons of consistency and clarity for Dutch readers, I will stick to the Dutch acronym.

(Stănescu 2018, 104). This shows how scientific “facts” about the genetic make-up of envisioned groups, however generalized, while appearing as objective and neutral information, again, do not exist in isolation of colonial and, sometimes extremist, political discourses.

Additionally, the biologically essentialist distinction between North-West Europe and the rest of the world resonates with the way in which different clusters of migration backgrounds were distinguished in Dutch governmental statistics and policy until 2021, between so-called “*Westerse*”, or Western, “*allochtonen*” and “*niet-Westerse*”, or “non-Western”, “*allochtonen*”. This distinction was officially based on the presence or absence of “economic and cultural similarities” between different groups with people from the Netherlands (see Keij 2000). However, the distinction between “Western” and “non-Western” has racial overtones which will be discussed further in Chapter 3 of this dissertation. The most obvious reminder of these racial overtones are the reasons for clustering people with particular migration backgrounds, mentioned by the sociologist Hilda Verweij-Jonker in her influential book *Allochtonen in Nederland*:

The discussed groups (1) have come to the Netherlands in large numbers and in a relatively short time, (2) have been subject to some form of government policy to facilitate their arrival in Dutch society, and (3) are clearly physically recognizable because of their skin color and a language incomprehensible for Dutch people. (Verweij-Jonker 1971, 7, translation by me)

The reference to both the “incomprehensible” language and the “clearly physically recognizable” skin colour of people with a migration background signifies that the word “allochtoon” and its subcategory “non-Western allochtoon” are more than merely “ethnic” categories. It is for the reasons mentioned above that I will not follow the governmental vocabulary in calling their classifications “ethnicities” or “migration backgrounds” in the period between 1970 and 2021, but rather opt for the combined word “race-ethnicity” (see also Yanow and Van der Haar 2013; Yanow, Van der Haar, and Völke 2016). In this light, the artwork of Van Breugel with its focus on milk as a defining cultural *and* biological characteristic of people in general and Dutch people in particular, can be read, too, as a race-ethnic construction of humanity.

With its scientific basis in contemporary state-of-the-art biological research and its attempt to represent an objective and universal account of humanity, the “Portrait of the Global Citizen” is reminiscent of Leonardo DaVinci’s well-known Vitruvian man. Ironically, through the inclusion of race-ethnic components, Van Breugel’s artwork is haunted by the very same issues as the Vitruvian man in its Eurocentric and colonial approach to the idea of the human (see Braidotti 2013, 13–16). Such Eurocentrism is “a structural element of our cultural practice” (Braidotti 2013, 15), which naturally finds its way into many of the technological apparatuses of science and governance. In the past, too, people

deemed biologically inferior, encompassed not just others who were visibly different but also others who were other not by virtue of their skin colour but by virtue of some other trait, now (newly) culturally loaded as designating inferiority. (Griffin and Braidotti 2001, 226)

As I will argue in this dissertation, the selection of such a “new” trait, to be “culturally loaded” with ideas about superiority and inferiority, is intimately tied to the techniques and technologies available to scientists and bureaucrats; racialization, understood as the “the extension of racial meaning to a previously racially unclassified relationship, social practice or group” (Omi and Winant 2015, 64), is inevitably mediated by technology in our contemporary society. As such, racializing practices of technologies are not only accidental byproducts or unconscious leftovers from premodern times, but are rather a structurally embedded element of the practices of modern science and governance. In the words of the sociologist and philosopher Zygmunt Bauman:

Racism is unthinkable without the advancement of modern science, modern technology, and modern forms of state power. As such, racism is strictly a modern product. Modernity made racism possible. (Bauman 2007, 61)

Therefore, if we are interested in the production and construction of race-ethnicity in contemporary practices of governance, we need to go beyond the understanding of racism as an interpersonal event, and investigate it as an institutional, discursive, and technologically mediated phenomenon.

When we consider modern race thinking not merely as a preexisting condition that works its way through, or is mediated by, new technologies, but simultaneously as an effect of those technologies, the role of datafication in the artwork of Van Breugel takes on a new meaning. While the visualizations of DNA in the portraits of the King and Queen clearly communicate the datafied construction of its information, it is no longer readable by humans. The “somewhat veiled, subtle form of transparency” (Van Breugel in Kleijn 2019, 28), provided by the abstract visualization, actively disconnects the presented information from the not so subtle and very problematic discourses it is haunted by. In this dissertation, I understand the disconnect caused by such practices of datafication as a form of “colonial aphasia” (Stoler 2016), understood as:

a condition in which the occlusion of knowledge is at once a dismembering of words from the objects to which they refer, a difficulty retrieving both the semantic and lexical components of vocabularies, a loss of access that may verge on active dissociation, a difficulty comprehending what is seen and spoken. (Stoler 2016, 12)

Only because of colonial aphasia, as the active dissociation caused by the abstract datafied presentation of the “facts”, can the artwork present a rather biologically essentialist construction of Dutchness in such a prominent location. I think it is safe to say that a more pictorial

presentation of all the problematic elements discussed above would have never been approved by the royal family, Dutch journalists, and the Dutch public. Therefore, the explicit presentation of particular parts of the genetic heritage of the royal family in a form that is unreadable for the average person, is one of the ways in which datafication obfuscates the ideological assumptions and social norms embedded in technologies.

## Research Questions

It could be argued that the datafied dissociation of the biologically essentialist discourses of Dutchness in the royal palace is merely the effect of an artwork. However, the underlying problem is that the technologies regularly used for practices of governance are connected to the very same discourse. As discussed above, the information infrastructure provided by CBS used categorizations with racist and colonial connotations. The data in which they were used were, unfortunately, not merely presented in brick on a wall, but rather used in practices of governance that had a direct and often detrimental effect on the lives of many Dutch people. In the period this research project was undertaken, between 2017 and 2022, the Netherlands was confronted with a series of scandals involving various Dutch governmental organizations using race-ethnic categories in their data systems resulting in instances of discrimination. First, there was the SyRI court case, in which the Dutch state was sued for the use of a fraud detection system called *Systeem Risico Indicatie*, which was using copious amounts of private data in order to create risk profiles mainly of people living in “problem areas” (see van Schendel 2019). Many scholars and activists were concerned about possible discrimination and unfair outcomes of the law and its accompanying algorithmic system (see Timan and Grommé 2019). It was eventually deemed too invasive in terms of privacy, lacking the proper legal grounds necessary for such an invasive method, and, therefore, abolished (see de Vries 2020).

A second case that attracted a lot of media attention in 2020, was the so-called *Toeslagenaffaire* (Child care benefits scandal), through which it became publicly known that the Dutch Tax and Customs Administration had illegally used data about second nationalities, migration backgrounds, and even skin colour of Dutch citizens in an algorithmic profiling system (see Autoriteit Persoonsgegevens 2020, 16). As a consequence of this practice, approximately 26,000 families, many of which had members with a migration background, were forced to pay back the subsidies they had received for their child care; subsidies they were legally entitled to in the first place, because their income was below a minimum threshold. Consequently, the individuals and families involved suffered great financial, mental, and physical distress, culminating in, for example, the removal of more than 1000 children from their families by the Dutch child protective services (see CBS 2021e). At the end of 2021, the Dutch government had still not paid families the money they were legally entitled to, and the full extent of their suffering had not been investigated yet.



A final case that recently received media attention is the *Rotterdamwet* (Rotterdam Act), a law which enabled Dutch municipalities to ban people with a criminal record, or without a job, from underprivileged neighbourhoods. This law now explicitly suggests that municipalities use the Leefbaarometer, an algorithmic data visualization which I will discuss in more detail in Chapter 4, which features an algorithm that uses non-Dutch migration backgrounds as negative indicators for livability, in its evaluation (see Uitermark, Hochstenbach, and van Gent 2017). Taken together, these examples show how contemporary forms of racialization involve various data technologies and algorithms in meaning-making processes that concern the migration backgrounds of Dutch people. Furthermore, these racializing practices go much further than mere stereotypical and biologically essentialist representations of people with or without a migration background; they have a direct effect on the living conditions and possibilities of all people living in the Netherlands.

While the childcare benefits scandal and the SyRI case were both based on the misuse of *private* information, the enforcement of the Rotterdam Act is not. The Leefbaarometer, together with the other case studies of Chapters 4 and 5, are systems that make use of *public* data—data that is openly available through the Application Programming Interface (API) of CBS. This interface allows third party software to directly tap into the aggregated information in the databases of the main provider of statistical information about Dutch society. This raises questions, not only about the ways in which data with and about race-ethnicity is used currently, but also about what the very existence of such data means in a society that is increasingly relying on data technologies for its day-to-day governance. If we want to prevent future scandals involving the use and misuse of information about the migration backgrounds of Dutch citizens and inhabitants, we need to better understand both the connection of race-ethnic categorization with histories of racism and colonialism and the ways in which the embeddedness of race-ethnicity in information infrastructures creates the conditions of possibility for discrimination. Therefore, the main question of this dissertation is:

***What are the practices and consequences of the datafication of race-ethnic categories in the Netherlands?***

I will divide this research question into three sub-questions, in order to be able to address different parts of this main question in more detail. First, I aim to discuss the history of contemporary practices of classification and show how they are not only socially defined, but, more precisely, *sociotechnically* constructed. By discussing the history of race-ethnic classification and its tools, we are able to better understand the histories and connections contemporary classification practices and instruments are haunted by. Second, race-ethnic classifications are not merely employed in bureaucratic registration practices, but are made available in statistics and databases through web based APIs. We need to address how race-ethnic information makes its way from the moment it is registered to the place where it is made available for statistical and algorithmic purposes. Third, the process of datafication creates

possibilities for the use of race-ethnic classification in any situation in which it is deemed effective by civil servants, programmers, activists, or journalists. I will, therefore, investigate three of such applications in more detail in Chapters 4 and 5. The sub-questions that will guide me in the aims mentioned above are:

1. What are the origins of the Dutch race-ethnic categorization system and how are its categories defined?;
2. How are race-ethnic categorization systems in the Netherlands datafied by government institutions?;
3. How are race-ethnically categorized data appropriated by governmental and non-governmental organizations?

With these questions, I aim to show how contemporary practices of race-ethnic classifications and their datafication are not only haunted by histories of racism, colonialism, and migration, but also take on new meanings through the technologies that have mediated them. By employing three different approaches to technologically mediated racialization, which I call the *instrumental*, *epistemological*, and *ontological* approach, I will show how data technologies do their work on several different levels. They are not merely instruments available to solve a problem, not merely epistemological tools that allow us to perceive and know the world around us, but ontological actors that co-shape our world and ourselves in that world. Datafied technologies that employ race-ethnic classification, therefore, do not merely use ideas about who counts as Dutch and who does not; they co-shape these very ideas.

## On Positionality

It should be mentioned that my position as a white Dutch researcher, who would have been classified as *autochtoon* in the period between 1970 and 2017, is not impartial. As offspring from the tulip farmer family Van Schie, who were living at the the *Zwarteweg*<sup>6</sup> in Noordwijk, and the family of Gladiolus cultivists named Van der Hulst from Warmond,<sup>7</sup> the phrase “pulled from the clay” could apply to me without much imagination. The etymology of the word Schie can even be traced back to the very Dutch practice of digging canals. Furthermore, my position as a white able-bodied, cisgender, straight man is also constructed by the technologies that not only inform and co-shape my identity, but are also explicitly made for me. As a result, I have to continuously

---

<sup>6</sup> This street name could be translated as Black Road, a name that has nothing to do with race, but rather with the original colour of the pavement (see A. H. Meijer 1992, 232).

<sup>7</sup> A popular family story is that of my grandfather discovering a new variety of Gladiola in his field which he chose to name after the local soccer club “Warmunda”, rather than after my grandmother (see <https://www.gladiolus.nl/catalogus/tubergenii/tubergenii-charming-lady>)

remind myself that the perspective that is culturally assigned to me is insufficient to lay bare the relation between race-ethnicity and governmental knowledge producing practices, since, in the words of feminist STS scholar Sandra Harding:

my eyes are not my own, nor are even my most private thoughts entirely private; they belong to my historical period-and to particular class, race, gender, and cultural commitments *that I do not question*. (Harding 1991, 100, italics in original)

Like the data technologies under scrutiny in this dissertation, my vision, language, and capabilities of understanding, are haunted by the information that is included in *and* excluded from my cultural context.



*Figure 2: Celebration of Sint Nikolaas at the Van Schie family, December 1988. The face of Black Pete is invisible due to a combination of very dark face paint and the photographic technology of the 1980's which was calibrated for white skin.*

Take for example the photo that was taken during de celebration of Sint Nikolaas at my grandmother's house in 1988 (see Figure 2). On the far right, it shows how the almost two-year old me is unpacking a present while sitting on dad's lap, while my five-month old sister is sitting on the lap of Sint Nikolaas in the centre of the picture. On the far left, the controversial figure of Black Pete is shown (see Helsloot 2012; Brienen 2014), but their face paint is too dark for the photographic technologies of the 1980s. One of the main reasons for this technological bias

towards white skin colour has to do with the so-called “Shirley cards”. These cards featured pictures of white women with several coloured squares or circles, which were used in the process of photo development to produce the “right” colours (Roth 2009, 112–15). This norm was only changed and expanded after manufacturers of chocolate and wooden furniture complained that they could not properly photograph the colour of their products in a realistic way (Roth 2009, 119–20). Thus, race as a cultural and technological blind spot becomes visible in this childhood picture in two ways: 1) this family scene that was and is often still not experienced as racist by many white Dutch people, is perceived as a cute opportunity for a photograph, and 2) the camera that was perfectly capable of photographing Sint Nikolaas, my dad, my sister, and me is, like many Dutch people, unable to properly perceive Black Pete. That is, as a cultural tradition that, like the epistemic technologies it is photographed by, is haunted by a past of colonialism and racism.

Only after spending a significant time abroad, most notably over a year in Turkey, I learned to see both the history and contemporary politics of the cultural Other in the Netherlands from a different perspective. Reading Gloria Wekker’s *White Innocence: Paradoxes of Colonialism and Race* and Rosi Braidotti’s *The Posthuman* back to back, while staying at Koç University in Istanbul in the summer of 2016, was instrumental, not only because I came to understand the problematic notion of objectivity and impartiality in the Western project of knowledge production, but also its implications for the treatment of the “allochtoon” Other in Dutch history, social sciences, and governance. These experiences provided me with the possibility to approach these social issues from another perspective, to be able to read the dominant Dutch grand narratives against the grain. In the words of feminist STS scholar Donna Haraway:

The knowing self is partial in all its guises, never finished, whole, simply there and original; it is always constructed and stitched together imperfectly, and therefore able to join with another, to see together without claiming to be another. (Haraway 1991, 193)

One of the most prominent places where my partial knowing self, as a recently graduated MA in New Media and Digital Culture, saw this problem playing out at the time, was in datafied tools like the “Allochtoon-o-meter”, a web-app providing the percentages of “allochtonen” in postcode areas to future home owners. This application, which I will discuss in more detail in Chapter 4, was made available on the alt-right weblog *Geenstijl* and made use of CBS data. The networked nature of the application made me question how much of the responsibility of the apparent racism embedded in the *Allochtoon-o-meter* could be attributed to CBS, which, after all, provided the necessary data. Another question that logically followed was for what other purposes race-ethnic data from CBS could be used. These considerations were the starting point for the investigation that led to this dissertation.

The reader should be aware that at the moment of writing my research proposal, between the summer of 2016 and the spring of 2017, I did not foresee that during the course of this

investigation all the labels of classifications would be changed, and the problematic clusters of migration backgrounds formerly termed “*Westers*” and “*niet-Westers*” would be abolished by the Dutch government, due to the stigmatizing connotations these word picked up over time. However, as I will argue throughout this dissertation, merely changing the labels of categories that are used in a datafied infrastructure does not affect the process of technologically mediated racialization.

## Chapters and Case Studies

In order to be able to properly address issues concerning race-ethnicity *and* issues concerning datafication, I will need appropriate theoretical tools. In Chapter 1 of this dissertation, I will therefore develop the concept of “technologically mediated racialization”, and explain the three different ways in which this concept can be understood and operationalized, each with its own benefits and drawbacks. First, I will argue that the ways in which ideas about race and ethnicity become implicated in the construction of Dutchness should be understood as a process of *racialization*. Within the fields of history and the social sciences, racialization, which can be explained as “the extension of racial meaning to a previously racially unclassified relationship, social practice, or group” (Omi and Winant 2015, 64), is usually not investigated in relation to the technologies that mediate this process. Therefore, I propose the concept of technologically mediated racialization to understand the ways in which the selection of “particular human characteristics for purposes of racial signification” (Omi and Winant 2015, 110) is not only historically and socially determined, but also in part the result of the specific technologies that aid in the process. I will continue to explain how, depending on the meaning of mediation, the phrase technologically mediated racialization can be understood in an instrumental, epistemological, and ontological way. The instrumental approach focuses mainly on the content of a system. It invites questions about the definitions of race that are used, and whether or not race is an “efficient” human characteristic to use within a given datafied policy situation. The epistemological approach invites us to problematize the ways in which knowledge is produced in systems. This allows for a discussion on how knowledge production is interwoven with culture, history, and ideology. However, broader questions about the ontological implications of the use of particular technologies for the process of racialization remain difficult to address. The final, ontological approach, therefore, asks questions about the relation of a data system to the world it is creating. Through this approach, we understand race in terms of a relational ontology that is dependent, not on biology or culture, but on the technologies that mediate its construction.

In Chapter 2, I will discuss how the three approaches to technologically mediated racialization can be operationalized in the investigation of the datafication of race-ethnicity in the Netherlands, by situating the approaches within the fields of *critical data studies* and *postcolonial studies*. The specific combination of these two fields provides me with the theoretical and methodological tools necessary for the investigation of how Dutchness and non-Dutchness are

constructed in the data infrastructures of the Dutch government. The three main methods that follow from critical data studies and postcolonial studies are *critical discourse analysis* (CDA), *affordance analysis*, and *infrastructural inversion*. With this combination of methods, we will be able to 1) critically assess the meanings and connotations of race-ethnic categories used in systems both in definition and in operation, 2) situate data systems in their context, reconstruct their perspective, and show how their technical possibilities afford the mediation of processes of racialization, and 3) trace the flows of data from data infrastructures to individual systems and vice versa, showing the institutional and infrastructural nature of technologically mediated racialization in the Netherlands.

The aim of Chapter 3 is to make visible how discourses on Dutchness and non-Dutchness have developed in the past century. Taking the centralization of the production of statistics about demographics in the Netherlands in the form of CBS as its starting point, I will trace how definitions concerning nativeness and foreignness developed from 1899 until 2020. Throughout this timeline, I will situate different categorization systems within their historical, social, and technical contexts. This means that I will take into account 1) historical events that have had an effect on the categorization systems, 2) social policies that have had an effect on what the Dutch government considered necessary to report on, and 3) the technical context of the practice of counting and computation. Through this historical approach, I pursue three main aims: 1) to situate current practices of race-ethnic categorization, 2) to denaturalize the current state of affairs concerning categorization, and 3) to show how categorization practices always should be considered within their technical context. The first aim relates to my argument that race-ethnic categorization as currently practiced, while largely obscured through changes in labeling, still relates to Dutch colonial history and migration history. In line with existing research, this notion goes against the persisting dominant idea present in Dutch nationalist discourses of colonialism as a thing of the past (see El-Tayeb 2011; Stoler 2016; Wekker 2016). My second aim is to show that the current state of affairs concerning technologically mediated racialization cannot be seen as a 'logical', 'natural', or 'normal' way of presenting the Dutch population. By showing the development of race-ethnic categories throughout the previous century, I aim to demystify how we currently look at categorization practices of the Dutch government and create a fresh start, before moving to contemporary datafied practices. My final aim, then, is to show how major shifts in race-ethnic categorization tend to happen in conjunction with the implementation of new technologies. It is important to note that I do not mean to say that one causes the other, as I am neither arguing in favour of social constructivism, nor of technological determinism concerning the relation between technologies and race-ethnic categories. Through my historical approach, however, I can show that major changes in either part of the relation tend to happen shortly after or in conjunction with major changes in the other. This historical approach will further strengthen my argument that in the investigation of racialization, the technological context should always be taken into account. I will end this chapter with a description of the

current state of affairs when it comes to race-ethnic categorization and the governmental data infrastructure of CBS that makes data available through an API (*Application Programming Interface*) creating possibilities for third parties to create their own applications that use data and statistics of CBS for their own purposes.

In Chapter 4, I will analyze two applications that make use of the CBS API, namely the Allochtonenmeter and the Leefbaarometer. Through a comparative approach, I aim to make visible the perspectives and politics of both these applications. In addition, I will reflect on how these perspectives and politics relate to the infrastructure on which they were built. The Allochtonenmeter is an application built on the CBS API, which, with a very simple design, returns the percentage of non-western immigrants for an entered four digit postal code<sup>8</sup>. The datafied nature of the application suggests that the practice of the Allochtonenmeter cannot be seen as an 'unintended' byproduct of an otherwise neutral epistemological infrastructure. Rather, racialization is an inherent process of an infrastructure that provides race-ethnically categorized data for re-use. Contrary to the independent, right-wing oriented Allochtonenmeter, The Leefbaarometer (LB) is an official governmental application which was ordered in 2007 by the Dutch Ministry of Internal Affairs. Where the Allochtonenmeter features visual parts that communicate its situatedness and politics, the Leefbaarometers presentation signifies objectivity and neutrality. However, when analyzed in terms of its affordances, I will show that it employs a similar positionality as the Allochtonenmeter, and that it too features forms of technologically mediated racialization. By comparing these apps, I aim to make three different, but related arguments. First, I will show that the combination of a discourse analysis and affordance analysis can enable a researcher to reconstruct the perspective and politics of a data technology, even if it is presented as politically neutral. The second argument I will make is that the perspective embedded in data technologies like the ones under scrutiny here tends to be the normative and culturally dominant perspective in a given context, which in the case of the Netherlands is a white, autochtoon adult with a job. When such a perspective is made explicit, it can hardly be argued that systems such as the Leefbaarometer benefit all Dutch citizens equally. On the contrary, these systems actively partake in practices of technologically mediated racialization. The third argument, then, is that the technicity for these racializing practices is already present in the infrastructure. Through an ontological approach, I will be able to argue that the Allochtonenmeter and Leefbaarometer are not accidental instances of technologically mediated racialization but rather materializations of a technicity of race that was already embedded in both the available technologies and related race-ethnic categories. The process of technologically

---

<sup>8</sup> [www.allochtonenmeter.nl](http://www.allochtonenmeter.nl)

mediated racialization does not work within particular technologies, but should be seen as an institutional and infrastructural phenomenon.

In Chapter 5, I will investigate the Crime Anticipation System (CAS) of the Dutch National Police as an example of a data technology that does no longer explicitly use race-ethnic categories. However, when investigated as a *data assemblage*, within its cultural and organizational context, it can be argued that CAS aids in the selection of human characteristics for racial signification. In the case of CAS, this is partly due to its design, and partly due to the way in which it was implemented and the specific organization it is implemented in. With its design that firstly targets only petty crimes, and secondly, uses characteristics such as family size, income, and proximity of known offenders, CAS does target specific neighbourhoods that will usually largely overlap with areas populated with people with a migration background. Combined with the track record of the Dutch police concerning issues of racism in the working environment and ethnic profiling (see Amnesty International 2013; Fassin 2013; Çankaya 2012; 2015; 2020; Gowricharn and Çankaya 2017), CAS becomes an assemblage that is particularly prone to technologically mediated racialization. In choosing CAS as a case study, I have threefold aims. First, CAS is yet another example of a system partly built on a CBS infrastructure, and while it does no longer feature race-ethnicity, it did use race-ethnic categories during its implementation phase. In addition, given the regular updates, and the rather “innocent” attitude of the Dutch when it comes to issues concerning race-ethnicity, there is no guarantee that race-ethnic categories will remain turned off in the system. Second, by explicitly including the organizational context in the analysis of CAS, it can be shown that race-ethnic categories are not the only thing to look for in the investigation of technologically mediated racialization. CAS as a data assemblage still contains racializing elements in the form of common racist practice by police officers and an organizational culture that largely resists attempts to diversify. A final argument that I will make, then, is that no matter which approach is chosen, whether it be ontological, epistemological or even instrumental, CAS should never have been implemented in 2017. It is one of the most obvious examples of technological solutionism in practice, an assumed self-fulfilling prophecy that still needs to prove itself.

In the concluding chapter, I will first answer the three sub-questions and the main question that guided this investigation. Then, I will discuss possibilities for future research that can be derived from the results of this investigation. In the final part of the conclusion, I will discuss how governments working with data systems, particular the Dutch government, can benefit from the results of this research. The current situation of data technologies used in governance practices, and especially the routine use of race-ethnic categories in them, calls for tools aiding the prevention of algorithmically generated racist outcomes, and guidelines for ethical and anti-racist use of categories in situations where they can benefit marginalized groups. I will show how the order of my analysis in this investigation of technologically mediated racialization—which goes



from instrumental use, through epistemological critique, to ontological deconstruction—can be inverted and used in the planning, evaluation, and design processes of current and future governmental practices that are aided by data technologies. In this way, I hope that this investigation is not only aiding in the production of knowledge within the scholarly fields of critical data studies and postcolonial studies, but also in guiding the Netherlands in the process towards a more accountable and just datafied society that values all people equally.

# Chapter 1: Understanding Technologically Mediated Racialization

In our contemporary datafied society, processes of inclusion and exclusion are no longer merely carried out by people amongst themselves, but increasingly involve interactions with various technologies (see for example Noble 2018; Eubanks 2018; Benjamin 2019). Many governments, the Dutch government included, now use algorithmic systems in their practices of governance. Systems used in these governmental practices often deal with information regarding race, ethnicity, or nationality. In the Netherlands, many such systems assign different numerical values to different countries of origin, which are then used to calculate crime risk, livability scores for neighbourhoods, real estate value, welfare fraud risk, and many other variables that are deemed necessary in the governance of Dutch neighbourhoods and cities (see van Schie and Oosterloo 2020; van Schie, Smit, and López Coombs 2020).

To adequately address these processes and conditions, we need conceptual tools that are appropriate for an investigation that is able to deal both with *social* issues related to race and *technical* issues concerning mediation and computation. I will argue that the process through which racializing connotations of geographic, ethnic or national ancestry are created or amplified should be understood as *technologically mediated racialization*. In the next paragraph, I will unpack this combination of terms. First, I will discuss how I understand *racialization* by engaging with postcolonial and critical race theory. I will argue that existing theorizations of racialization barely mention the mediating function of technology, something that is essential in the investigation of racialization in contemporary governmental data systems. Second, I will discuss three different takes on *mediation* that largely shape research on biases and inequalities in datafied processes. Each of these approaches implies a different view on how data technologies and society relate to each other, as well as a different understanding of the mediating role of datafied technologies in the process of racialization.

The three approaches to technologically mediated racialization that I recognize are the *instrumental*, *epistemological*, and *ontological* approach. Since I will discuss these approaches in much more detail in the middle part of this chapter, I will only briefly introduce them here. Within the instrumental approach, data technologies are portrayed as neutral and objective instruments that can be used to make governance processes more efficient and more effective. This approach—which is very prominent in data science, business, and governance discourses—advances a rather positivist and empiricist idea about the relationship between the world and knowledge production. In short, when systems make use of categorizations such as race and ethnicity they tend to take them as a fact of life, rather than as a local, constructed, and situated way of producing knowledge (Haraway 1988). The second approach, i.e. the epistemological

approach to technologically mediated racialization, tries to unpack those assumptions embedded in datafied systems by taking into account the inherently political nature of knowledge production. This way of thinking—prominent in critical data studies and science and technology discourses—takes technologies as epistemic actors, each of which offers specific possibilities for racialization. By employing this type of social constructivist frame, some critical digital race scholars for instance blame biases on human factors, such as skewed historical data, racist programmers, or biased users. Unfair outcomes of systems are then often framed as mistakes or byproducts instead of features inherent to a particular technologically mediated practice. In order to be able to reflect on the features of such inherently racializing technologically mediated practices, I suggest that we explicitly add a third approach to our conceptual tools: an ontological approach to technologically mediated racialization. This approach does not assume that racialization precedes its technological mediation, but argues instead that it is entangled in the process; it both shapes mediation and is shaped by mediation. Within this way of thinking—prominent in philosophy of technology discourses (see, for example, Ihde 2009; Verbeek 2012)—data technologies mediate our lifeworld and produce us as subjects in that lifeworld.

While important research had been carried out within the paradigms of each of these approaches at the start of this investigation in 2017, most of this research was interested in either racialization and discrimination *or* in data and algorithms. I will argue that the small but increasing number of studies that is interested in the combination of the two, tend to work within instrumental and epistemological understandings of datafication. The aim of this dissertation is to add an ontological approach to technologically mediated racialization to the available conceptual and theoretical options of both critical data studies and critical race studies. This would enable a more comprehensive understanding of how current knowledge producing processes in governmental data systems are entangled with the ways in which our lifeworld is shaped. Such an understanding aids in assessing the politics of race-ethnic classification in governmental data system; it aids in assessing whether such classifications constitute forms of technologically mediated racialization or whether they aid in affirming societal inequities. In the remainder of this chapter, I will first discuss the concept of racialization and then discuss in more detail the aforementioned instrumental, epistemological, and ontological approaches to technologically mediated racialization.

## **Racialization**

After a meeting in 1967, UNESCO issued an important statement that thoroughly explained the fallacies implied by a biological understanding of different races considered to have existed continuously for centuries (see UNESCO 1969). This statement, written by the most prominent scholars on the topic at the time, stated that race should no longer be understood as a physiological characteristic of people which can be known through the investigation of skin color, body shape, and ancestry (Hiernaux 1969). Instead, a social constructivist approach was

proposed in which the meaning of race is considered a context-dependent social reality (see Banton 1969). Through such an approach, questions concerning what race *is* no longer make much sense; instead, the focus shifts to the *process* through which race gets its meaning and shapes reality. This process is commonly called *racialization*. Racialization should not be confused with racism. Where racialization designates a meaning making process, racism signifies an ideology that informs negative attitudes and behavior towards people deemed of a different race (Todorov 1993). In this paragraph, I will first discuss racialization as a context-dependent discursive process of othering based on ancestry, nationality and/or ethnicity. In the final part of this paragraph, I will address how this understanding of racialization holds up in the specific context of the Netherlands.

Racialization happens in the interplay of diverse elements such as language, institutions, and epistemic practices that together produce meanings about particular concepts, objects, and people. These knowledge producing systems invested in one particular topic or area – such as race – is what Foucault calls a discursive formation (Hall 1997a, 44–45). The totality of such a system, often referred to as the *dispositif*, or apparatus, does not just produce knowledge or truth, but creates the very conditions for “the separation, not of the true from the false, but of what may from what may not be characterized as scientific” (Foucault 1980, 198). The politics of this *episteme*, in its historical and geographical setting, produces a situated *régime of truth*:

that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements, the means by which each is sanctioned; the techniques and procedures accorded value in the acquisition of truth; the status of those who are charged with saying what counts as true. (Foucault 1980, 131)

This system, that is made up by institutions, technologies, and people, produces the very conditions that are necessary to decide what counts as knowledge and truth, and what does not. As such, régimes of truth are the starting point for many studies dealing with institutionalized knowledge production about the inhabitants of a country, organized and legitimated by state actors.

Régimes of truth are intricately linked with the ways in which a state relates to its inhabitants. In his earlier work, Foucault shows how régimes of truth work and produce specific subject positions in the context of the prison system (Foucault 1977), psychiatry (Foucault 2006), and sexuality (Foucault 1978b; 1988; 1978a). In later works, however, Foucault moves away from power/knowledge on the level of the individual and starts to investigate how the state exerts power over its citizens by considering the population as a whole. In his lecture series titled “Security, Territory, Population”, Foucault (2007) discusses the gradual development in which European countries moved from rule based on sovereignty, largely based on force and obedience, towards an *art of governance*, based on the creation of conditions in which

populations act as desired, out of their own volition. Within this framework, “power is situated and exercised at the level of life, the species, the race, and the large-scale phenomena of population” (Foucault 2013, 43). Such a “biopolitics” and “governmentality” (the accompanying philosophy of governance) are informative concepts in the investigation into the relationship between a government and its population, and the normative implications of that relationship. Governmentality in this context can be understood as

[T]he ensemble formed by institutions, procedures, analyses and reflections, calculations, and tactics that allow the exercise of this very specific, albeit very complex, power that has the population as its target, political economy as its major form of knowledge, and apparatuses of security as its essential technical instrument. (Foucault 2007, 144)

The process of defining the population that is subject to governance, and the desired behaviour of that population often includes the production of normative ideas about race and gender (Stoler 1995; 2002; Rasmussen 2011; C. Taylor 2011). Former colonial and continuing colonial power relations can be traced within such specific normative assumptions that are the result of biopolitical measures, structures of governmentality, and their legitimation through regimes of truth.

In feminist and postcolonial studies, the discursive process of associating negative meanings with non-normative racial, ethnic, or gender identities is understood as “Othering” (Said 1979; Spivak 1988; Hall 1997b; Bhabha 2010). “Othering” is how discriminatory meaning and practice are invented, justified, and perpetuated. Building on Saussurian semiotics and poststructuralism, Stuart Hall explains how in this meaning making process, the relation between a sign, called the “signifier”, and the concept it refers to, called the “signified”, only emerges through “binary oppositions like Us/them” (Hall 1997b, 258). Referring to Jacques Derrida, he reminds us that “we are not dealing with (...) peaceful coexistence (...) but rather with a violent hierarchy. One of the two terms governs (...) the other or has the upper hand” (Derrida 1981, 41, as cited in Hall 1997b, 258). In each of these hierarchical binaries, like Man/woman, White/black, Autochtoon/allochtoon, the term that signifies the societal norm becomes the default term, while the non-normative term becomes the Other. This process can happen to a wide variety of human characteristics and axes of difference, ranging from sex, race, and ability, to gender, ethnicity, and class. The specific forms of difference under scrutiny in this dissertation are those human differences that can be understood as being connected to historical forms of racialization based on ancestry, broadly known under the labels “race”, “ethnicity” and “nationality”.

Building on the work of Foucault, Edward Said (1979) conceptualized the discursive fabrication of the Other in the colonizer/colonized relationship “Orientalism”. In his influential work of the same title, Said set out to explain how Orientalism manifests itself in science, cultural representations, and politics. Produced by cultural representations and legitimated by science,

Othring informs power hierarchies both between nations as well as between people within a nation. As Said explains:

Debates today about Frenchness and Englishness in France and Britain respectively, or about Islam in countries like Egypt and Pakistan, are part of the same interpretive process, which involves the identities of different "others," whether they be outsiders and refugees, or apostates and infidels. It should be obvious in all cases that these processes are not mental exercises but urgent social contests involving such concrete political issues as immigration laws, the legislation of personal conduct, the constitution of orthodoxy, the legitimization of violence and/or insurrection, the character and content of education, and the direction of foreign policy, which very often has to do with the designation of official enemies. (Said 1979, 332)

One of the main reasons for this process of Othring is the fact that the production of culture and knowledge is predominantly performed through a white, Western, and male perspective (Harding 1991; Puwar 2004). Groups of people that do not fit this norm have historically been excluded from the spaces in which knowledge production and governance take place (Puwar 2004). As a result, in the knowledge that is produced they are subjected to processes of Othring, which manifest in various forms of exclusion, as well as in the formation of stereotypes (Puwar 2004, 142–43). Critical theorist Homi Bhabha (2010, 110) explains that “[l]ike the mirror phase ‘the fullness’ of the stereotype—its image as identity—is always threatened by ‘lack,’” meaning that both the colonizer and the colonized envision the other in terms of difference. As I will show throughout this dissertation, governmental data systems take the perspective of the colonizer, in the form of the white “*autochtoon*”, as the default position and envision the discursively constructed Other only through a perceived lack, based on geographic, ethnic, or racial ancestry. I understand this Othring, based on geographic, religious, ethnic or racial origins, through the concept of “racialization.”

Racialization is understood in this dissertation as “the extension of racial meaning to a previously racially unclassified relationship, social practice or group” (Omi and Winant 2015, 64). As a process, racialization is continually producing racial formations, which I understand as geographically and historically situated collections of people, things and practices, which are grouped on the basis of their perceived race, ethnicity, or nationality. Such an understanding contests both essentialist views the concept of race as something objective, biological, and concrete, and social constructivist views on race as an “illusion” born in social relations and discourse (Omi and Winant 2015, 109). According to historians Michael Omi and Howard Winant we should, therefore, understand race as a concept

*...that signifies and symbolizes social conflicts and interests by referring to different types of human bodies. Although the concept of race invokes seemingly biologically based human characteristics (so-called phenotypes), selection of these particular human features for*

purposes of racial signification is always and necessarily a social and historical process. Indeed, the categories employed to differentiate among human beings along racial lines reveal themselves, upon serious examination, to be at best imprecise, and at worst completely arbitrary. They may be arbitrary, but they are not meaningless. Race is strategic; race does ideological and political work. (Omi and Winant 2015, 110–11, italics in original)

It is precisely in governmental bureaucracy that the ideological and political nature of categorizations is most prominent but is, at the same time, denied in the process of presenting classifications as a scientifically neutral process. While I fully subscribe to the thesis that the selection of human features for racial signification is socially and historically situated, what is missing from this definition is the role that different technologies play in these very processes. Therefore, in this dissertation I will argue that racialization should not merely be understood as a social and historical process, but also as a technologically mediated process.

A pervasive argument in ongoing debates in the Netherlands is that race and racialization have been conceptualized primarily to investigate racism in an Anglo-Saxon context and would, therefore, not be suitable for an investigation in the Dutch context. The main reason for this argument is that in the dominant Dutch race discourse, “*ras*” is still discursively tied to the genocide of Jewish people and the eugenic practices of Nazi Germany (Yanow, Van der Haar, and Völke 2016, 190). When use of the word became prohibited in policy contexts in the Netherlands after WW II, its discursive meaning remained relatively frozen until recently. As a result, “*ras*” is, still, in daily Dutch discourse primarily understood in a very narrow biological sense. To avoid using the term, Dutch researchers have been relying on concepts such as ethnicization (see Koyuncu-Lorasdağı 2013; Bonjour and Block 2016) or minoritization (Rath 1991). This is in sharp contrast with the development of race discourse in the United States and the United Kingdom, where “*race*” is used to convey cultural and political meanings (Rath 1991, 117–18). Here, race used to be understood in a biological sense as well (see Hiernaux 1969), but ongoing societal discussions, the civil rights movement, and the more recent Black Lives Matter movement have altered the discourse in such a way that the word “*race*” is now a concept that is available to address societal injustices. As a result of the incongruency between the Dutch and Anglo-Saxon discourses, to speak about current affairs in terms of “*race*”, and by extension “*racialization*” in the Netherlands, is deemed to be anachronistic and out of place (see Bovenkerk 1984; Vuijsje 1986, and more recently the report by the Scientific Council for Government Policy by R. P. W. Jennissen et al. 2018). There are, however, several objections to be made against this reasoning.

First, it is impossible to draw hard lines between the concepts race, ethnicity, and nationality in bureaucratic practice. Empirical policy research on governmental classifications has shown that nationality, race, and ethnicity are often used interchangeably or in conjunction with each other. In the Netherlands, this has even led to classification systems used in perinatal care and obstetrics which feature nationality, ethnicity, and racial classifications within the same system, meaning

that people have to choose “the most appropriate” category of overlapping classifications such as “Caucasian”, “Turkish”, “African”, “Negro” and “Hindustani” (see Yanow, Van der Haar, and Völke 2016). Furthermore, until 2021, Statistics Netherlands (*Centraal Bureau voor de Statistiek*, from now on CBS) used a classification system – which I will discuss at length in Chapter 3 – based on what they call “ethnicity”, which is technically based on the formal place of birth of someone’s parents. This explanation is not complete, though. One example of the way in which the “ethnic” categories are racialized here can be found in the statistics on “ethnic diversity” in elementary schools; at least until 2015, government reports have labelled schools in which the majority of students have a migration background as “black schools” (see de Mooij et al. 2015, 11). This confusion and/or interchangeability has led scholars to adopt hybrid terms such as “ethnonationalism” (Connor 1994), “ethnoracial” (Hollinger 2005), and “race-ethnicity” (Yanow 2003) to address the fuzzy nature of nationality, ethnicity, and race related concepts. The work of policy researcher Dvora Yanow is especially relevant in this context; she investigated bureaucratic categories in both the United States and the Netherlands with a similar conceptual apparatus and concluded that categories in both countries have ethnic and racial meanings (Yanow 2003; Yanow and Van der Haar 2013). Yanow and Van der Haar argue that, even if the Dutch category of “allochtoon” is technically primarily based on ethnic and national concepts, the way it is operationalized in Dutch governmental practice is undoubtedly racialized (Van der Haar and Yanow 2011). It is for this reason that they understand the Dutch categorization practice as “race-ethnic”. This argument is in line with a growing body of research that addresses the racializing nature of Dutch migration and integration discourses (Essed 1991; Wekker and Lutz 2001; Ghorashi 2014; Essed and Hoving 2014b; Özdil 2014; Wekker 2016). Therefore, I assume that the concept race is not anachronistic at all and that it does play a significant role in Dutch policy and bureaucracy. Therefore, it deserves proper attention.

A second objection against the supposed misfit of the concept racialization in the Dutch context, is that racialization is particularly suitable to move the discussion from the essence of categories to the way in which they operate. As I will elaborate further in this chapter’s section about mediation and technology, the question is not only whether labels of the categories deployed carry inherent racial meanings, but even more importantly how those labels are operationalized. For example, it is rather easy for the Dutch government to deny that race is playing a role in government policy, given that there are no explicit race categories in demographic data. However, through the lens of racialization, it is possible to investigate the *process* through which negative meanings are attached to certain words, situations, or visible characteristics of human bodies. This means that the Dutch labels “allochtoon” and “autochtoon” not only carry racial meanings to begin with, the manner in which they are operationalized racialize the categories they represent even further. As I will show in Chapters 4 and 5, “autochtoon” tends to function as the default category in many systems, while “allochtoon” is made into the Othered category in these same systems partly because it is consistently connected to negative human



characteristics, such as crime and poverty. One of the arguments I will develop in Chapter 3 is that the government-sanctioned name change of the label “allochtoon” to “person with a migration background” in 2016 and 2017 did nothing to change its operationalization in data systems, and by extension the racializing character of the category. In short, racialization as a process may happen to identity characteristics that, in principle, have nothing to do with skin colour, ancestry, or other characteristics with historical racialized discriminatory and exclusionary practices. Therefore, the official absence of racial categories does not mean that racialization does not take place. On the contrary, as the process of racialization is invisibilized, it becomes much harder to spot and, therefore, a more disguised and insidious process.

Finally, denying that race may play a role in Dutch society and that, consequently, racialization is an anachronistic and out-of-place concept could itself be a symptom of the way in which the Dutch national/ethnic/racial identity is constructed as tolerant and non-racist. In Dutch academic discourse, traces of this dominant way of thinking can be found in the early work of Dutch cultural anthropologist Frank Bovenkerk (1978b, 7), who, in 1978, concluded that while racial discrimination does exist in the Netherlands, it is “not as ubiquitous as in the United States, since the [Dutch] ideology resists the phenomenon more unequivocally” (Bovenkerk 1978a, 193–94, translated by the author). This assumption of exceptionalism from Bovenkerk is contradicted—in the same book—by the statement of John M. Alsswang, an American professor teaching at Leiden University at the time, who claimed that despite the perceived tolerance towards minorities, the Dutch stance on racial discrimination was covered by “a well-intentioned conspiracy of silence” (Alsswang as quoted in Bovenkerk 1978b, 21). The most well-known theorization of this “conspiracy of silence” has been produced by Gloria Wekker, who calls the Dutch attitude “white innocence”, which she explains as

[The] dominant way in which the Dutch think of themselves, as being a small, but just, ethical nation; color-blind, thus free of racism; as being inherently on the moral and ethical high ground, thus a guiding light to other folks and nations. (Wekker 2016, 2)

Wekker argues that this self-image causes white Dutch people to experience moments of cognitive dissonance when they are confronted with their own racist remarks and behaviour. It is this attitude combined with the feeling that contemporary societal structures are completely unrelated to the Dutch colonial past, that makes it particularly difficult to address institutional racism in the Netherlands.

It has often been argued that this apparent disconnect between colonial history and contemporary societal structures should not be seen as accidental, but rather as an inherent feature of the need of white colonial nations to match their history with their self-image of rationality, equality, and progress (Stoler 2016; Wekker 2016). Additionally, this self-image coincides with ideas of homogeneity, which makes the mere presence of minorities into “an

intolerable deficit in the purity of the national whole” (Appadurai 2006, 53). Derrida (1982) characterized this attitude as a “white mythology”; as the ways in which Western metaphysics “erased within itself the fabulous scene that has produced it” (Derrida 1982, 213). However, in contradiction to the official histories, the past “remains active and stirring” and is in a way “inscribed in white ink” (Derrida 1982, 213); present but only visible when actively looking for it, and always capable of being “reactivated” (Stoler 2016). One possible explanation for this “white mythology” (see also Young 2004) in the Dutch context has been offered by Zihni Özdil (2014), who argues that the apparent amnesia concerning the colonial past of the Netherlands should be understood in terms of a “pasteurization” of history. The concept of pasteurization, usually referring to the process by which milk is treated so “it will not upset the consumer’s stomach” (Özdil 2014, 50), metaphorically connects the significance of milk in the Dutch cultural archive (see introduction), with the process through which Dutch history—as it is taught in schools and museums—is made easier to digest. By leaving out the parts that are too shocking or politically inconvenient, such as slavery and colonialism (see Weiner 2014a), the collective memory of the Dutch is severely hampered.

Several scholars have argued that amnesia is too mild of a term for this phenomenon, opting for the use of the more profound “*cultural aphasia*” (Helsloot 2012) or “*colonial aphasia*” (Stoler 2009; 2016; Wekker 2016). As discussed in the introduction of this dissertation, the historian Ann Laura Stoler explains *aphasia* as an occlusion of knowledge:

It is not a matter of ignorance or absence. Aphasia is a dismembering, a difficulty in speaking, a difficulty in generating a vocabulary that associates appropriate words and concepts to appropriate things. Aphasia in its many forms describes a difficulty in retrieving both conceptual and lexical vocabularies and, most important, a difficulty in comprehending what is spoken. (Stoler 2016, 128; emphasis in original)

Aphasia in this sense is not the absence of colonialism from collective memory and institutions, but the inability to access these memories through language. According to Wekker (2016, 2), this has caused the Dutch difficulty in addressing the existing “deep structure of inequality in thought and affect based on race”. She argues that it is from this “deep reservoir, the cultural archive, that, among other things, a sense of self has been formed and fabricated” (Wekker 2016, 2). Also referring to Stoler, Wekker explains the cultural archive as a “repository of memory” (Stoler 2009, 49), of which the contents are both “in the heads and hearts of people”, but also “silently cemented in policies, in organizational rules, in popular and sexual cultures, and in commonsense everyday knowledge” (Wekker 2016, 19). The aforementioned name change of the label “allochtoon” to “person with a migration background” without changing the definition and operationalization of the category would be a perfect example of colonial aphasia, since it actively dismembers and silences the meaning of racializing organizational rules in Dutch policy, while keeping its functionality in place.

For all of the reasons mentioned above, I do not consider racialization as an anachronistic or out-of-place concept. Rather, in this dissertation I argue that it is the most fitting way to address how policy, bureaucracy, and by extension, data systems used in a Dutch governmental context can attach racial meanings to categories through their functioning. However, to properly address how the processes of racialization and colonial aphasia happen in data infrastructures and algorithmic systems, we need to expand racialization theory to account for these specific technological contexts. In the next section, I will, therefore, theoretically explore different ways in which we can understand what I call *technologically mediated racialization*.

## **Technologically Mediated Racialization**

As argued in the previous section, racialization does not only happen in social processes between groups of people, but increasingly also involves various data technologies and algorithmic systems. I, therefore, understand the process of racialization as it happens in data technologies as *technologically mediated racialization*. While the terms “technologically mediated” might seem rather self-explanatory in this context, they can be understood in at least three different ways which should not be confused with each other. Drawing on the work of media scholar Joshua Meyrowitz (1999), I will discuss these three different ways, and extend them into three related approaches to the investigation of technological mediation: an instrumental approach, an epistemological approach, and an ontological approach. In conjunction with these approaches, I will discuss recent critical race studies and postcolonial studies on data and algorithms and argue that most of these studies tend to work within the epistemological understanding of technological mediation, while studies that work within the ontological understanding are generally scarce. The aim of this dissertation, then, is to fill this gap and create possibilities for adding analytical depth to future research, as well as to be a guide for and in future policy making.

### **The Instrumental Approach – Racialization as Content**

The first approach to technological mediation follows a rather common understanding of how people interact with tools and instruments: people pick the most suitable tool for a situation at hand and then use it to solve a problem, produce knowledge, or otherwise get a task done. In this understanding, the problem that needs to be solved exists independently of the available tools to solve it, and the available tools are there to help people do their tasks. In such an anthropological definition of technology (Heidegger 1977b, 5), instruments exist to serve people rather than the other way around. They are mere facilitators that are in no way related to the issues at hand, or the way in which we view the world. Like hammers are made to drive nails in wood, data technologies are made to aid us in the production of knowledge. When we take this lens and look at media technologies, we see them as instruments that are able to transport content to another time or place. In this case, the “medium” can either be understood as the system or technology which is transporting a message, such as television or internet, or as the

carrier of this information, such as a vinyl record or a USB memory stick. This content can then be used, read, or viewed – by another person or system, ideally in the exact same way as intended by the maker. Media technologies which are used as aids in knowledge production are envisioned to do just that: provide an objective and unaltered account of reality, but with an added bonus, usually in the form of increased detail or speed.

The instrumental paradigm largely follows the *media-as-vessel/conduit* metaphor of media as explained by Joshua Meyrowitz, in which mediation is a rather value free and neutral process that happens *between* two or more actors. In this, by now rather dated understanding, the content of media is analytically separated from the technologies that carry, send, and/or receive it (Meyrowitz 1999, 45). Mediation, by definition, takes place in the middle, in the “in-between, *as such*” (van den Eede 2011, 140). In this process, technology is but a mere facilitator, a tool the effects of which depend on the intentions and proficiency of its users. Through this approach, media technologies are understood as instrumental; they can be used well or badly, but do not necessarily embed values or politics themselves. Knowledge producing media, such as data technologies, are usually chosen for tasks *because* of their in-betweenness, as in practice, the transparency of their functionalities and the opacity of their embedded social values are often mistaken for neutrality and objectivity once their designs have stabilized to some extent (Feenberg 1999, 96–97; van den Eede 2011, 153).

The perceived neutrality of the process of mediation as performed by data technologies is clearly visible in the early popular and professional discourse surrounding data technologies. Two by now infamous articles published on Wired.com can be seen as exemplary of a general techno-positive, overconfident and rather uncritical attitude towards the possibilities of data technologies that has been convincing many businesses and governments. First, in the article titled “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”, Chris Anderson (2008) argued that the best way of understanding information is to look at it “mathematically first and establish a context for it later”, as, according to him, massive amounts of data can replace many of the existing sciences and philosophies:

Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves. (C. Anderson 2008, n.p.)

In a similar fashion, in an article titled “Big Data and the Death of the Theorist”, Ian Steadman (2013) suggests that

[i]n the same way that the internal combustion engine spelled the end of the horse as a working animal, big data could be the tool to render host of academic disciplines redundant if it proves better at building better narratives of human society. (Steadman 2013, n.p.)

The rapid adoption of data technologies as a new transparent tool to render or mediate reality quickly gained traction amongst companies and a diverse range of governmental agencies at the local and national levels (see Mayer-Schönberger and Cukier 2013; Kitchin 2014b; Wohlers and Bernier 2015; Schäfer and van Es 2017). This instrumental paradigm has become popular in The Netherlands (see Kool, Timmer, and van Est 2015; WRR 2016, 49–64). In response, critical scholars have started to speak about the condition of our contemporary society in terms of a “datafied society” and argue for a “critical understanding of this current situation and its societal consequences” as well as an attempt to “debunk the exceptionalism inherent in the ‘big data’ paradigm” (van Es and Schäfer 2017, 13).

“Big Data’s” perceived exceptionalism holds that many social and societal issues are solvable through the use of data to either make existing bureaucratic and managerial tasks more efficient and effective by taking humans out of the process in particular situations. Through this way of working, any societal problem is made into an engineering problem that can be solved with a technical solution. However, such a data-centric “technological solutionism” (Morozov 2013) runs the risk of obfuscating possible solutions that do not incorporate computational technologies. Due to the pervasive idea that any problem can be solved by applying the right datafied method, possible solutions that are more hands-on, or merely based on non-digital policy, are often overlooked.

Furthermore, in an instrumentalist perspective on technologically mediated racialization, data about identity characteristics such as ethnicity, race, or nationality, can also quite literally become instrumental. The selection, measurement, and processing of peoples’ characteristics is seen as a neutral endeavour, as instrumentalism generally also favours a rather empiricist understanding of knowledge production. The assumption here is that even in matters of ethnicity or race, there are objective standards, which can be known when the right tools are used. In this worldview, knowledge about skin colour, countries of origin, and religion can be recorded as facts and from that moment onwards be used to cross-reference with other demographic and statistical characteristics of people. This is not generally seen as a political or cultural matter, but rather as an objective and neutral scientific endeavour.

However, when this datafied process of knowledge production is compared with the aforementioned process of racialization, the similarities are striking. The abovementioned description of datafied empiricism basically explains “the extension of racial meaning to a previously racially unclassified relationship, social practice or group” (Omi and Winant 2015, 64) and the selection of “particular human features for purposes of racial signification” (Omi and Winant 2015, 110), in somewhat more neutral and scientific sounding terms. I, therefore, argue that an instrumental approach to technology invariably leads to instrumentalism regarding race.

Rather than being an accident, racialization becomes a feature, an add-on or plug-in, which can be turned on when programmers deem the use of this information “effective”.

In a context where information about race, ethnicity, and nationality is readily available, instrumentalism makes governmental data projects prone to the intentional and unintentional implementation of racializing features. In such contexts, race quite literally starts to function *as technology* (Chun 2009; Coleman 2009). Its availability in the Dutch context can be understood as a testament to instrumentalism as the main mode in which both technological *and* race-ethnic matters are considered. In addition, the availability of race-ethnic information, such as the availability of data technologies, seemingly justifies its use. One of the programmers of the Leefbaarometer I spoke with during the investigation for Chapter 4, remarked that *not* using ethnic information in the creation of the Leefbaarometer would have been a political decision, and he and his colleagues “did not engage in politics”. This shows that the default way of thinking about both data technologies and race-ethnic data is that using them in automated policy decisions is 1) a logical step and 2) an objective and neutral act, while the suggestion of diverging from this frame is seen as a political move. Thus, an instrumentalist view makes clear demarcations between what is considered scientific and objective and what is political and, therefore, subjective and vice versa.

The inconsistency of this attitude becomes clearly visible when people attempt to use race-ethnic categorization in non-racializing ways, for example in an attempt to increase inclusivity and diversity in settings where it is still largely lacking. When the Dutch Parliament had to vote about implementing the registration of the migration backgrounds of students and faculty/staff in Dutch higher education as an inventory before setting targets in terms of diversity and inclusivity, the proposal was dismissed. The motion filed against the proposal by the MPs Wiersma, El Yassini, and Van der Molen, which stated that people’s identities “should not be reduced to their ethnicity or migration background”, received a majority of the votes and was implemented.<sup>9</sup> Such remarks are, however, rarely heard when matters of crime, fraud, or welfare are discussed. As I will show in the case studies in Chapters 3, 4 and 5, race-ethnic information is routinely considered relevant in these situations, and objections such as the ones mentioned above are rarely expressed by politicians in relation to issues such as crime or fraud detection. Thus, when it comes to the production of knowledge about social issues in the context of government policy, the scientific and political always necessarily merge. The choice to gather data and produce knowledge *in a particular way* is inherently ideological.

---

<sup>9</sup> This quote comes from motion nr. 230 in case nr. 29338 (VSO Nationaal actieplan voor meer diversiteit en inclusie in het hoger onderwijs en onderzoek) discussed on 25 November 2020 by the Dutch house of representatives.

Race-ethnic information can function as a means to an end in particular situations, and can even be beneficial when used in a critical and evaluative manner, and when the envisioned ends—such as inclusivity in higher education—require this knowledge. In situations such as these, the political goals could temporarily trump epistemic precision, a practice referred to in feminist and postcolonial discourses as “strategic essentialism” (Eide 2016). In the words of postcolonial scholar Gayatri Spivak: “You pick up the universal that will give you the power to fight against the other side, and what you are throwing away by doing that is your theoretical purity” (Spivak in Spivak and Harasym 2014, 12). Such a pragmatic use of essentialism should, however, be treated with care, and only temporarily, to prevent the reification and naturalization of categories one eventually wants to get rid of. My criticism of instrumentalism in policy decisions should, therefore, not be read as a complete dismissal of all forms of ethnic, racial, or religious classifications in data systems, but rather as a call for a contextualized, situated, and critical use of data science. Using racial categories can be racializing in one situation and work in anti-racist ways in another; the absence of race is sometimes desirable while in other contexts it would constitute a less favourable form of “colourblindness” that attempts to transcend “racial labeling and the associated implicit and explicit biases to a hypothetical and almost mythical post-racial stance” (Braddock 2021). To be able to distinguish between the above mentioned situations, we need approaches that go beyond a merely instrumental consideration of technologically mediated racialization. We need an approach that is able to take into account the ideological and epistemological aspects of knowledge production as well as their relation. In the next section, I will discuss this type of epistemological approach to technologically mediated racialization.

### **The Epistemological Approach – Racialization as Operation**

The second approach to technologically mediated racialization, which is very common in STS and critical data studies discourses, is increasingly becoming part of both popular and professional debates. Comparable to the work done within the instrumentalist paradigm, this epistemic approach separates content from the technologies that mediate its storage, distribution, and consumption. However, where in the instrumentalist approach media are considered neutral transporters of messages, the epistemological approach focuses on the logic within which the content is produced, stored, and distributed. The epistemological approach, then, draws attention to the range of possibilities media have to offer; it looks at what messages *can* be mediated as well as the specific ways in which they *are* mediated. As such, mediation is, therefore, not a passive transference of a message, but rather an active shaping of that message. The ways in which media shape messages depends on both industry conventions and on the equipment that is used. By taking a look at medium specific characteristics of individual technologies, we can analyze their implied power structures:

While a content analyst exploring women's "images" in media may be concerned with the roles held by the women (e.g., housewife vs. executive), a media grammar analyst might

examine whether women in a film - regardless of role - are framed in intimate space, made to look weak through high angle shots, or sexualized by voyeuristic shots of their body parts. (Meyrowitz 1999, 48)

The framing of women in films can be seen as one way in which content is shaped by industry conventions. However, when we extend Joshua Meyrowitz's explanation from audio visual media to data technologies, and understand mediation in terms of datafication, we are able to not only look at how human characteristics such as race, ethnicity, and nationality are represented in systems as a result of historically ingrained societal conventions, but also how the selection of "particular human features for purposes of racial signification" (Omi and Winant 2015, 110), is shaped by the material options different data technologies have to offer. With an epistemological lens we are able to look at processes of racialization as a form of politics that is the result both of institutional conventions *and* the design of specific data technologies.

The politics of technologies in processes of knowledge production has been one of the core subjects in the field of science and technology studies (STS). The word "politics" is meant here as a designation of the "political causes and consequences of technical artifacts" (Brown 2015, 6), and should not be confused with the more common use of the word politics to describe the activities of politicians. STS scholars such as Bruno Latour, Sandra Harding, and Donna Haraway have all examined the socially constructed and non-neutral nature of knowledge production. Sociologist Bruno Latour's (2005) actor-network theory (ANT) aids in the investigation of the production of knowledge as a process that emerges from sociotechnological assemblages including both human and non-human actors. Such dispositifs, or networks, consist of interrelational materialistic and semiotic elements of objects, subjects, and actors/actants (Latour 2005). Within Latour's vocabulary, data systems cannot be interpreted as neutral intermediates that collect, calculate, and visualize statistical data without steering and influencing the transference of knowledge embedded in these processes. Data functions as an "immutable mobile", a displacement of knowledge "through transformation" (see Latour 2005, 223, note 308). Such a transformation, also called a "translation" (Latour 1991) necessarily changes the possibilities particular information has to offer. The translation of information about ethnicity, race, or nationality into data increases the scope of that content as well as broadening its applicability in automated policy contexts. The translation of the physical and digital world into data has, therefore, always ideological implications, some of which can be predicted and some of which will only emerge after implementing a system or technology. Within ANT, a translation is an explicitly non-neutral form of mediation that invites an investigation into its workings, not to invalidate its ways of producing knowledge, but rather to make visible its politics and make possible an informed discussion about the implications of such a mediation.

Feminist critiques of scientific objectivity can be particularly helpful in addressing the politics and implications of knowledge producing assemblages such as data systems. Feminist STS scholar



Sandra Harding distinguishes two possible modes of critique in the fight against epistemic inequalities in her influential work *Whose Science? Whose Knowledge? Thinking from Women's Lives* (1991). The first approach, which she calls “feminist empiricism”, sees social biases in the outcomes of scientific research as bad scientific practice (Harding 1991, 111–18). When prejudice ends up in scientific work, this is rarely the result of structural or institutional issues, but can usually be traced back to human conduct. This perspective on “biased” knowledge production implies that neutral and objective knowledge is possible if researchers adhere sufficiently strictly to their own standards of objectivity. The benefit of this approach in eliminating inequalities in science is that it does not challenge the basic structures of science as an institution, and that it adheres to traditional views on what counts as reliable knowledge. This is, however, also its downside, as instances of “bad science” are generally seen as coincidental and unrelated, rather than as the result of structural issues.

The second approach Harding recognizes and advocates, uses a different understanding of how knowledge relates to the world. Such “feminist standpoint epistemology” suggests that actors and agents involved in knowledge production practices should be attentive to power relations, and should incorporate in their work who benefits from a particular perspective, and who does not (Harding 1991, 119–37). With feminist standpoint epistemology, a normative perception can be made visible which acknowledges what is left unseen or what is excluded in a knowledge producing assemblage. Objectivity, in this view, does not so much lie in the information that is extracted from the world, but rather in the way in which researchers communicate the perspectives that are used in their research. To acknowledge one’s position, as, in my case, a white, European, middle class, heterosexual, cisgender male, is seen as a form of “strong objectivity” (Harding 1991, 149), i.e. much better than pretending that I somehow produce knowledge in a neutral and impartial way. The latter practice is rather common in the natural and much of the social sciences, where people tend not to make visible in their texts and reports who produced the knowledge that is presented. Such a working practice would be understood in feminist standpoint theory as a form of “weak objectivity”(Harding 1991, 143). As I will show in the case studies of this dissertation, data systems also rarely explicitly communicate their situatedness and would, therefore, also fall in the category of “weak objectivity”. However, data systems do generate knowledge from a particular vantage point, and feminist standpoint theory enables us to recognize that perspective.

Moreover, feminist standpoint epistemology does not only criticize the construction of objectivity prevalent in much of the natural and social sciences, but also makes a value judgment about which perspectives are most useful in laying bare the power structures that guide and shape the production of knowledge. Harding argues that outsiders to power structures are particularly capable of making them visible:

Using women's lives as grounds to criticize the dominant knowledge claims, which have been based primarily in the lives of men in the dominant races, classes, and cultures, can decrease the partialities and distortions in the picture of nature and social life provided by the natural and social sciences. (Harding 1991, 121)

While the statement above talks about women's lives in particular, it would work just as well for culturally disadvantaged people in any other axis of difference. Where I, as an able-bodied person, would generally not be aware of the one or two steps up necessary to enter most buildings of Utrecht University, any person in a wheelchair would be able to tell in an instant which buildings would need accessibility improvements. Similarly, as a white person, I would never directly find out whether or not a particular facial recognition system is biased against people with darker skin tones. In Heideggerian terms, infrastructures, systems, and technologies which are *ready-to-hand* for me, might be *present-at-hand* for people with different physical and non-physical identity characteristics. Technologies being present-at-hand are at best a nuisance and at worst dangerous in daily life; in research, however, situations of breakdown, or miscommunications between people from marginalized communities and technologies, present opportunities in the investigation of technologically embedded cultural and social inequalities. Harding's standpoint theory, therefore, assumes that people in an oppressed situation have the most direct experience of discriminatory social structures, and are most capable of producing knowledge about the instruments, infrastructures, and social values that shape their oppression.

The standpoints of the 'subjugated' are, however, not innocent either. The prominent feminist STS scholar Donna Haraway argues for critical reflexivity in the production of what she calls "situated knowledges" (Haraway 1991, 191). Its aim should go beyond the production of a variety of standpoints in order to show a kind of relativism, as

[r]elativism is the perfect mirror twin of totalization in the ideologies of objectivity; both deny the stakes in location, embodiment, and partial perspective; both make it impossible to see well. Relativism and totalization are both 'god-tricks' promising vision from everywhere and nowhere equally and fully, common myths in rhetorics surrounding Science. But it is precisely in the politics and epistemology of partial perspectives that the possibility of sustained, rational, objective enquiry rests. (Haraway 1991, 191)

Haraway's often used phrase of the "god-trick" points at a particularly extreme form of "weak objectivity", a form of disembodied knowledge production that historically shaped much of the natural and social sciences. Instead, researchers should aim to provide a "partial perspective" that is accountable and clear about its position and politics.

This idea of the "partial perspective" can also be useful in assigning a perspective to existing epistemic practices. When all perspectives are partial, the knowledge produced by technology, for example a data system, is necessarily partial as well. The theory of situated knowledges can then be used in the reconstruction of the often obfuscated perspective of a data system. While I

will discuss this in more detail in Chapter 2 on methodology, for now it suffices to say that by taking a close look at the inputs, possibilities, visualizations, and outputs of a system, it is possible to reconstruct a perspective in its various stages of construction, as well as its embedded normative assumptions. Such an evaluation does not only include what is present in a system, but also what is missing. Since both norms and perspectives are often not made explicit, the gaze of these systems seems to come from nowhere in particular. Haraway aptly describes this perspective as

the gaze that mythically inscribes all the marked bodies, that makes the unmarked category claim the power to see and not be seen, to represent while escaping representation. This gaze signifies the unmarked positions of Man and White, one of the many nasty tones of the word objectivity to feminist ears in scientific and technological, late industrial, militarized, racist and male dominant societies. (Haraway 1991, 188)

The unmarked category in any power relation is the category that is seen as the norm, and as I will show in Chapters 3, 4, and 5, it is this category that is usually missing in data systems as well. When a system takes ethnicity into consideration, the people considered natively Dutch are not represented, while people with a migration background are. When income is considered as a variable in a system, the categories used in calculations usually contain percentages of people on welfare, but percentages of employed people are not always part of the equation. Such inconsistencies help both in reconstructing the assumed default position in systems, as well as situating the system's ways of producing knowledge. In addition, instead of getting stuck in pointing fingers at racist programmers and oblivious bureaucrats, we can focus on the process of knowledge production and in doing so uncover structural and institutional discriminatory and racializing practices.

One clear non-digital example of such an institutional racializing practice has been provided by media scholar Lorna Roth (2009), who showed that the stereotypical portrayal of black people in analog film and photography up until the early nineties had not only been the result of choices made by filmmakers and photographers. One of the major reasons for the lack of detail in the visual portrayal of people with darker skin tones was caused both by the material limits of film and cameras—which were optimized for lighter skin tones—at the time, as well as the industry standards in the development process of film. One of the main reasons for this technologically embedded predisposition towards white skin color has to do with the so-called “Shirley cards”. These cards featured pictures of white women with several squares or circles containing what were deemed the most important colours, and were used in the process of photo development to produce the “right” colours (Roth 2009, 112–15). This norm was only changed and expanded after manufacturers of chocolate and wooden furniture complained about the fact that they could not photograph the colour of their products in a realistic way (Roth 2009, 119–20). Only after some changes were made in camera technology as well as to the way in which “white

balancing” was done in film and photography were the circumstances created for a more detailed and nuanced portrayal of black people in visual media. This example shows how institutional racism can be embedded not only in the practices of people that produce messages, but also in the material, equipment, and infrastructure that is available to them.

Like analog film and photo equipment, a variety of digital technologies have built-in prejudices and biases against people with darker skin tones. Facial recognition systems and biometrics are technologies known in particular for their biases against black people (see for example Browne 2015; Cheney-Lippold 2017; Buolamwini and Gebru 2018; Wevers 2018). While these biases are mainly caused by the training data sets consisting of people overwhelmingly white and male (see also Buolamwini and Gebru 2018), the disparities between the ability of systems to recognize white and non-white people also point at a more structural and institutional shortcoming. Simone Browne theorizes these biases as “digital epidermalization”, caused in part by an embedded “prototypical whiteness” (Browne 2015, 26). Epidermalization, a term coined by the psychiatrist and decolonial thinker Frantz Fanon, deals with the ways in which race is imposed on bodies from the outside. In the words of Fanon himself:

I am overdetermined from the outside. I am a slave not to the “idea” others have of me, but to my appearance. I arrive slowly in the world; sudden emergences are no longer my habit. I crawl along. The white gaze, the only valid one, is already dissecting me. I am *fixed*. Once their microtomes are sharpened, the Whites objectively cut sections of my reality. I have been betrayed. I sense, I see in this white gaze that it’s the arrival not of a new man, but of a new type of man, a new species. (Fanon 2008, 95)

These examples should not be seen as a reversal of Haraway’s unmarked category claiming “the power to see and not be seen, to represent while escaping representation”. The visibility of normative identities in facial recognition systems and their absence in data systems are part of the same power structures in knowledge producing infrastructures. Similar to the aforementioned example of not wanting to “reduce” students and professors to their ethnicity, while crime and welfare are routinely made known in terms of racial and ethnic differences, in these cases, visibility, hypervisibility, *and* invisibility tend to work in favour of culturally dominant groups of people. After all, it is in the power of those groups to decide what stands out and what does not, or, in the case of the collection of data, what needs to be made visible and what does not.

In the context of racialization, visibility and invisibility should not merely be understood in terms of what information can be tied to human bodies, or what can be ‘seen’ by people. Increasingly, many identity characteristics available in data which are used to cross-reference with social phenomena such as crime, fraud, and welfare are invisible to the human eye and/or not necessarily about bodies. An epistemological approach makes it possible to investigate racialization beyond classically known notions of race in terms of skin colour or ancestry. Media

scholar John Cheney Lippold (2016) explains how the National Security Agency (NSA) of the USA could not legally use their surveillance systems on US citizens, only on non-Americans. However, electronic communications are hard to pin down to a specific nationality in real time. The NSA, therefore, devised a system that could algorithmically estimate the likelihood of people being American citizens or foreigners. This system presented a completely new take on the age-old designations *jus soli*—signifying a system in which one’s nationality is tied to the country of birth—and *jus sanguinis*—signifying a system in which one’s nationality is tied to blood relations. Cheney-Lippold called this new situation, in which the likelihood of being a ‘foreigner’ only needed to be 51% in order to be subjected to surveillance, *jus algoritmi*. It is no surprise that the NSA system disproportionately targeted people with family in other countries, and that this group largely overlaps with people who are already marginalized to begin with.

Another example of a system in which information is no longer tied to the bodies of people is a much debated predictive policing system implemented in the Dutch city of Roermond.<sup>10</sup> This system, aimed at decreasing shoplifting in the city center, notifies the police when small cars with license plates from eastern European countries cross the German-Dutch border and drive into the shopping center. Eastern Europeans, responsible for only 17% of shoplifting cases, are deemed the most important culprits in this datafied assemblage.<sup>11</sup> Both the license plates in Roermond and the electronic communications crossing the US border function as a “human feature” made available “for purposes of racial signification” - as the digital variant to the age-old stereotype. These examples show how the epistemological approach allows us to look beyond classically known practices of racialization. Rather than looking merely at systems that literally use race-ethnic categories, its aim is to bring forward how racialization as a process happens within particular systems, and between systems and the infrastructure within which they operate. The particular parts of a data assemblage that are of interest, then, are those parts which aid in the selection, transmission, and use of data in the process of racial signification.

One common misunderstanding that regularly makes its way into the analysis of racial biases in data assemblages, is the confusion of instrumental and epistemological approaches due to the fact that critiques of the knowledge producing capacities of systems still tend to focus exclusively on the human actors—instead of taking into account both human *and* non-human actors—in an assemblage. Such a covert instrumentalism can sometimes be seen in recent work on racism and discrimination in data systems by Safiya Noble and Ruha Benjamin. In her influential study on “algorithmic oppression” in the search algorithm of Google, Safiya Noble states that

---

<sup>10</sup> <https://www.oneworld.nl/lezen/discriminatie/hoe-nederland-a-i-inzet-voor-etnisch-profileren/>

<sup>11</sup> <https://www.amnesty.nl/wordt-vervolgd/criminaliteit-voorspellen-roermond-politie>

[p]art of the challenge of understanding algorithmic oppression is to understand that mathematical formulations to drive automated decisions are made by human beings. While we often think of terms such as “big data” and “algorithms” as being benign, neutral or objective, they are anything but. The people who make these decisions hold all types of values, many of which openly promote racism, sexism, and false notions of meritocracy, which is well documented in studies of Silicon Valley and other tech corridors. (Noble 2018, 1)

At worst this understanding of “algorithmic oppression” assumes the media, in this case data systems, are perpetuating human biases, void of any agency or politics of its own. At best, these technologies have politics, but these are explicitly the result of the scientific paradigm, social situation, or culture in which they are embedded. In her 2019 book on what she calls “the new Jim code”<sup>12</sup>, Ruha Benjamin often employs a similar line of thought:

Racism, in this way, is a technology that is “built into the industry”. But how does racism “get inside” and operate through new forms of technology? To the extent that machine learning relies on large “naturally occurring” datasets that are rife with racial (and economic and gendered) biases, the raw data that robots are using to learn and make decisions about the world reflect deeply ingrained cultural prejudices and structural hierarchies. (Benjamin 2019, 59)

In this way, any politics embedded in technology are always thought of as mistakes or foul play from humans rather than an inherent feature of particular technologies. The culprits of embedded inequalities are usually considered to be either biased historical data, biased designers, or biased users of a system. In such an analysis, the medium functions as a mere conduit; racism goes in, racism comes out, data systems are let off the hook. This is how instrumentalism creeps into what should be an epistemological analysis of the politics and ideological implications of both the data, and the systems that operationalize that data.

Benjamin’s modus operandi combines STS and critical race studies in order to “open wide” the black box of algorithmic systems (2019, 41). Like Noble, she reads the interaction between culture and the “surfaces” of the systems she tries to open up:

Race critical code studies, as I develop it here, is defined not just by what we study but also by *how* we analyze, questioning our own assumptions about what is deemed high theory versus pop culture, academic versus activist, evidence versus anecdote. The point is not

---

<sup>12</sup> A clever reference to the infamous anti-black Jim Crow laws that enforced racial segregation in the United States in the late 19<sup>th</sup> and early 20<sup>th</sup> centuries.

just to look beneath the surface in order to find connections between these categories, but to pay closer attention to the surfaces themselves. (Benjamin 2019, 45)

While the work of both Safiya Noble and Ruha Benjamin is very important in showing the continuities of racism in technologies that are in popular discourse often considered to be neutral, they do not scrutinize the particularly technological parts beyond the surfaces of the racist or abolitionist systems that are at the center of their investigations. Inherent to this approach is the risk that racial bias is theorized as a form of “bad science”, rather than a particularly racializing knowledge producing system that fails to account for its partial perspective. My aim is, therefore, to look at the inner workings of the data systems under scrutiny here, in order to properly situate them, before I question their implications. By adding the possibility for technological explanations to available cultural and social theory, I will be able to have a more technically informed discussion about the embedded politics of systems and how they relate to history and culture.

The understandable and justifiable reasons for not providing technological explanations in the arguments of Noble and Benjamin are plenty. First of all, many of the systems discussed are proprietary, meaning that the code is considered intellectual property and kept secret by their owners. Furthermore, even if the code is available, one would need the skills to read and interpret it, something that is often referred to as “data literacy” (see Gray, Gerlitz, and Bounegru 2018). This is a more general problem for humanities scholars and sociologists, who often lack the technical understanding to include the technicalities of data systems in their analysis. The epistemological approach enables one to understand media in the analysis of stereotypes, biases, and racializing tendencies in the operation of particular media. Like instrumentalism, an epistemological approach can produce valuable information about racialization, but it should never be the end point of any analysis. The epistemological approach aids in addressing the medium specific possibilities of data systems that mediate racialization, even, or especially, when this was not intended by their makers or users. In Chapter 2 of this dissertation, I will discuss in more detail how we can understand these possibilities through the concept of affordances. This concept makes it possible to investigate governmental open data systems by employing a combination of critical data studies and postcolonial studies lenses to the specific data systems in question.

A final limitation of the epistemological approach is that it merely looks at data systems as knowledge producing entities, where increasingly these systems engage more directly with our everyday lives. In the case of automated systems, the reflexive step where knowledge is interpreted and choices are made is often delegated to computers as well. The idea of computer logic bleeding into our lifeworld, therefore, invites us to think about technologically mediated racialization in more than an epistemological way. Data systems do not only report on our world, they may also enable and perform it. In addition, data systems do not produce all knowledge

themselves, they imply larger material and immaterial infrastructures that support it. Data systems can not be separated from our society, but are woven into the fabric of our lifeworld. To understand how this aspect affects racialization, we need to be able to go beyond epistemology and ask ontological questions. In the next paragraph, I will therefore introduce the possibility to think about technological mediation as an ontological condition.

### **The Ontological Approach – Towards a Technicity of Race**

A third and least commonly known way to understand mediation is to not think of it as something that happens between pre-existing actors, but as the source of emergence for these actors and the function they have in the relationship. In short, media do not only *communicate* reality to us, but could themselves be understood as a shaping force of our reality. This understanding of media as an environment themselves changes the focus from both the content and materiality of media towards the larger structures they imply, i.e. how media create the conditions for particular societies to emerge *and* how particular societies give rise to particular media that fit its structures well.

“Media determine our situation”, said German media scholar Friedrich Kittler, and, “in spite or because of it-deserves a description” (Kittler 1999, xxxix). Media scholar John Durham Peters (2015) explains how in Kittler’s understanding, media are “world-enabling infrastructures; not passive vessels for content, but ontological shifters”, and, as a result, “[i]nconspicuous vehicular transformations can have gigantic historical effects” (Peters 2015, 25). Thus, the ontological approach implies again a different understanding of mediation. Technologies are no longer understood as mere tools or knowledge producing actors, but as actively changing our lifeworld; technologies work on an ontological level. Mediation has become an ontological condition; its produces both our lifeworld and us as subjects in that lifeworld (see Ihde 2009; Verbeek 2012). Investigating *technologically mediated racialization* through this lens means asking questions about how race comes into being or how it is performed in relation to technology.

Within the ontological approach, we also need to rethink the meaning of technology. Technology is often explained as something along the lines of “the use of rational knowledge – whether it be scientific or technical – to satisfy needs, wants, desires or phantasies through the industrial creation, distribution, and management of goods and services” (Salomon 1984, 128). However, this definition tends to instrumentalize the technological, as if it were a neutral tool that objectively creates, manages, or produces. Instead, as the philosopher of technology Martin Heidegger explains, technology should not be seen as creating or producing, but rather as a *revealing*, since it makes us experience the world in new ways; it makes the world and its things appear differently (Heidegger 1977a, 13). In the case of modern technology, this revealing develops as a *challenging-forth* “which puts to nature the unreasonable demand that it supply energy that can be extracted and stored as such” (Heidegger 1977a, 14). This challenging-forth



does not only make nature, energy, and food available as resources, but also reveals humans themselves as resources within a particular technological rationality. It is here that the boundaries between technology and knowledge production are blurred, something that can clearly be read from the shared origins of the words *technē* and *epistēmē*, which both signify a particular type of revealing (Heidegger 1977a, 13). Through this ontological framework, we can understand data technologies as 1) materially embodied—as computers or machines, or *in* computers or machines, 2) employed in the gathering, storage, production, and dissemination of knowledge—in the form of data, code, algorithms, visualizations, and other interfaces that facilitate communication with both humans and other computers, and 3) embedded in organizational and governmental practices and infrastructures.

The question I aim to foreground with the ontological approach is how technological development drives processes of racialization. I will argue not only that technologies can carry racial politics, but rather that racialization as a process can be understood as *inherently technological*. Within such a framework, technologically mediated racialization and its product, racial formations, inherently find their origins in technologies aiding life, science, and governance. The project at hand, then, is to research and theorize how this process works by not only investigating the racisms that go in and out of systems, but also how specific racializations are produced through technological mediation.

To prevent our interest in the material and infrastructural aspects of media and technology to become uncritically teleological, I suggest that we use the concept of *technicity*. With this concept we can simultaneously account for technological as well as societal forces in the invention and development of racializing technologies. Heidegger has explained technicity in the form of the essence of technology which “is by no means anything technological” (Heidegger 1977a, 4). At first sight, this sentence might seem rather cryptic. After all, what does Heidegger mean with “essence” and isn’t the very idea of technology that it is technical in nature? The answer lies in his distinction between technology and technicity, between actual technological artifacts and their functioning, and the conditions of possibility for technology or technological thinking. Take for instance Heidegger’s example of the hydroelectric dam and its relation to the Rhine river. When does the river change from being a mass of water and a geographical phenomenon that shapes the surrounding landscape, into a source of energy? For Heidegger this does not happen at the moment the construction of the dam is finished. After all, the transformation of kinetic energy of the water into electrical power was not just a lucky byproduct discovered the moment the hydroelectric generator was turned on, but a process that was deliberately built into the design of the technology. Therefore, the transformation of the river does not happen because of technology and its functioning, but rather at the moment of the conception of the idea to build it, combined with the technical capacity to do so. This moment of conception transforms, or in Heidegger’s words “*reveals*”, the river as a source of electrical power. It is the technological

conditions of possibility that make the river available as a source of energy. In other words, the hydroelectric dam does not have to be built in order for the river to be available as a source of electric power. It is technicity, as rationality and as conditions of possibility, that makes the Rhine available as, what Heidegger calls *Bestand*, as a resource that can be *challenged forth*. The essence of technology should, therefore, not just be seen as technological, but as possibilities embedded in technical, social, and cultural relations.

When we apply the logic of the hydroelectric dam to data systems employing ethnic categories, it becomes clear that ethnic categories reveal people in racializing ways. People themselves become a *Bestand* available to be *challenged forth* in a racialized manner. The availability of ethnic classifications in open data of the government is, therefore, in itself a political and ethical matter, even before it is used in data systems. Through classifications such as the Dutch “allochtoon” and “autochtoon”, race functions as technicity, available to be used in systems for whatever purpose people think it is needed. Race as technicity bridges the gap between what is considered to be cultural and what is considered to be technical, because both culture and technology could be considered “activities of manipulation and thus techniques” (Simondon 2015, 18). However, where techniques are specific, systematized practices, and technologies are the materialized embodiments of these practices in the form of material objects, technicity as a concept is located at the level of the conditions of possibility for the development of individual technologies in an infrastructure. Technicity, again, should never be confused with technology, since the latter points at *actual* technical objects rather than the *rationality* and *conditions* from which they are derived. Just like a particular habitat is capable of housing particular organisms, technicity makes possible particular materializations in the form of actual technologies *as well as the people and societies that are living with those technologies*. In the context of technologically mediated racialization, race functions as a technicity; the situated racial formations produced by datafied technologies create specific social, technological, and cultural possibilities. In a datafied society, such technologically and infrastructurally situated racial formations are a condition of possibility for racializing technologies with detrimental effects for racialized people.

The existence of particular conditions of possibility for racializing technologies in the Dutch situation does not mean that all possibilities will eventually become a reality. After all, the development process of new technologies is not teleological in nature, and the direction of the development of particular materialization is indeterminate. In the vocabulary of Simondon, we would say that the evolution of technological objects tends to develop from abstract possibilities and ideas to concrete embodied technologies. The different parts of which a technology consists become increasingly integrated into a whole. Where in the beginning, different parts could fulfill functions in different machines, in later developmental stages this becomes increasingly less possible. Parts of machines become increasingly specialized and *concrete*; a process which Simondon calls *concretization* (Simondon 2017, 38–42). At the apex of development, the

technology is experienced as an *individual*, an entity that cannot be divided in parts, for the elements have integrated to the extent that they can no longer be considered parts. In such an *individual*, the parts can no longer fulfill any other functions, and the object, or ensemble of elements, has become one in its functioning in relation to its environment:

The compatibility of elements in a technical individual presupposes the associated milieu: the technical individual must therefore be imagined, which is to say presupposed as already being constructed in the form of an ensemble of ordered technical schemas; the individual is a stable system of the technicities of elements organized as an ensemble. (Simondon 2017, 74)

The concretization of elements and the individuation of ensembles always happens in relation to a particular environment, the “associated milieu”. It is important to realize that the associated milieu both gives rise to a new technology, but is also shaped by this new technology. Take for example the Leefbaarometer, an algorithmic system that we will explore further in Chapter 4. This system was built to measure livability based on CBS data, including information about the inhabitants of particular streets and neighbourhoods. Through the implementation of the Rotterdam Act (see Uitermark, Hochstenbach, and van Gent 2017), the Leefbaarometer now partly determines who is allowed to live in particular places. Not only are the results of this system, therefore, literally shaped by, and shaping the environments and people it is quantifying, it is also itself the result of the existing governmental rationality, *and* shaping the future of that governmental rationality. Technicity as a concept is, therefore, pointing at the mutual construction of the conditions of possibility of individual technologies, as well as their related environments.

Furthermore, when we take another look at the conceptual pair of *autochtoon* and *allochtoon* through the lens of individuation, we could think of these categories as tied to the technologies and structures in relation to which they were produced. The idea of a technicity of race, therefore, helps in situating particular racial formations and particular technologies as inherently connected through mediation. What we end up with, then, is a relational ontology of race in which racial formations are always mediated by specific data technologies; each technology produces a racial ontology based on its historical and sociotechnical context *and* technological possibilities. It should be noted that this stance does not equate to an ontologization of race—that is, considering race a reality that is merely made legible by technologies—but rather makes race into an object that does not exist in its own right; race needs actors, data systems in the case of this dissertation, connecting and producing its meanings and materiality. In the words of anthropologist and critical race scholar Amade M’charek: “race is not a singular object ‘out there’ in nature, but a relational entity enacted ‘in here’” (M’charek 2013, 421). The “in here” pertains to the scientific apparatuses and practice that mediate, and, therefore, continue to construct, process, and disseminate, race. In this context, the once hopeful proposition made by Paul Gilroy,

that “new ways of seeing, understanding and relating to ourselves point to the possibility that the time of ‘race’ may be coming to a close” (Gilroy 1998, 840), seems somewhat premature and naive in its belief that a change in technological mediation would somehow reveal race as the fiction it was all along. It does, however, as highlighted by Thao Phan and Scott Wark (2021), point at the necessity for an investigation into how the ontology of race is coupled with the instruments that mediate it, which in a datafied society equates to investigating “racial formation as data formations”:

If categories of race are inextricable from the technologies of classifying and sorting that makes the production of distinctions between people possible, it follows that technological innovations engender innovative ways of producing and policing difference. (Phan and Wark 2021, 4)

As researchers, we can use this understanding to construct ideas not only about how technologically mediated racialization happens now, but also about technicities of race—that is, the conditions of possibility for future racializations—currently present in society. In this way it might be possible to have political discussions about systems before they are implemented, rather than after the discovery of automated racializing and racist conduct of government agencies. In the creation of data systems we need some vision, therefore, of what we would like the future to look like. When systems are no longer merely treated as knowledge producing actors, but rather as world making agents we can be more aware of the fact that in the design process of a system racial categories should only be used when the ontologies produced aid in creating a more equal and just society. This way equity can function as a starting point, a guiding principle, in the production of data systems and algorithms, rather than as an afterthought.

## Conclusion

In this chapter, I proposed three different approaches to thinking about racialization in technological contexts. Each of the three approaches invites us to ask different types of questions (see Figure 3). The instrumental approach focuses mainly on the content of a system. It invites questions about the definitions of race that are used, and whether or not race is an “efficient” human characteristic to use within a given datafied policy situation. This approach usually does not question the ideological assumptions present in policy and practices. The epistemological approach invites to ask questions about ways in which knowledge is produced in systems. How are particular human characteristics selected for racial signification? And what politics are embedded in the process of racial signification in these particular systems? With these questions we are capable of discussing how knowledge production is interwoven with culture, history, and ideology. However, broader questions about the ontological implications of the use of particular technologies for the process of racialization are equally important to address. The final, ontological approach, therefore, asks questions about the relation of a data system to the world it is creating. Which racial formations are envisioned by a system and are created in the process?

What infrastructures is a system part of, how does a system itself function as an infrastructure, and finally, what technicities of race are consequently created? In answering these questions, we need to attend both to the data technologies that mediate processes of racialization as well as the historical, cultural, and organizational contexts in which they are designed and are operating.

<b>Approach</b>	<b>Types of research questions</b>
<i>Instrumental</i>	<ul style="list-style-type: none"> <li>• Which definitions of race-ethnicity are used in a data system?</li> <li>• Which societal issues are related to race-ethnicity?</li> <li>• Is it necessary or “effective” to connect a particular societal issue to race-ethnicity?</li> </ul>
<i>Epistemological</i>	<ul style="list-style-type: none"> <li>• Which human characteristics are selected for racial signification and how?</li> <li>• How is the selection of human characteristics for racial signification related to the material and informational characteristics of technologies?</li> </ul>
<i>Ontological</i>	<ul style="list-style-type: none"> <li>• How do racial ontologies develop in relation to technological systems?</li> <li>• What infrastructures are envisioned and created when a system is implemented?</li> <li>• Which technicities of race are created by infrastructures?</li> </ul>

Figure 3: Three approaches to technologically mediated racialization and their related research questions

Investigating the datafication of race-ethnic categories in the Dutch context through the three approaches to technologically mediated racialization will provide a broad understanding of the implications of the presence of race-ethnic categories in governmental data infrastructures. First, it allows us to recognize the types of reasoning that drive policy makers, researchers, and journalists when they discuss and critique data systems and practices of race-ethnic categorization. Second, placing comments and arguments within the instrumental, epistemological, and ontological approach allows for a recognition of the limitations and shortcomings of these arguments. Finally, these approaches could be used in the process of implementing new technologies and categories for the purposes of governance, by addressing each of the questions they raise and by checking whether or not their answers align with each other and with the broader intended policy goals.

In the next chapter, I will discuss the research fields that form the methodological foundation of this study, critical data studies and postcolonial studies. Furthermore, I will discuss how the three approaches of this chapter can be operationalized and employed within three methods common in these research fields in order to successfully study technologically mediated racialization in the Dutch governmental data infrastructure.

## Chapter 2: Investigating Technologically Mediated Racialization in Data Systems

Investigating the process of racialization in data infrastructures and data systems on an instrumental, epistemological, and ontological level requires multiple methodological tools. To properly account for the technological and social side of datafication and racialization, I will, therefore, combine methods from the scholarly fields *critical data studies* and *postcolonial studies*. It is important to note, however, that these scholarly traditions do not only provide methodological tools, but also provide particular critical ways of looking at media technologies, their cultural context, and the way in which they are entwined. The specific methods that I will use for my investigation, namely *critical discourse analysis*, *affordance analysis*, and *infrastructural inversion*, are methods particularly suited for the investigation of the datafication of race-ethnicity in the context of Dutch governmental data systems. With the described combination of methods, we will be able to 1) critically assess the meanings and connotations of race-ethnic categories used in systems both in definition and in operation, 2) situate data systems in their context, reconstruct their perspective, and show how they mediate processes of racialization, and 3) trace the flows of data from data infrastructures to individual systems and vice versa, showing the institutional and infrastructural nature of technologically mediated racialization in the Netherlands. In what follows, I will first discuss the two fields of study in which my work is situated and what I aim to contribute to them. I will subsequently discuss the methods I will be using in the three case studies of this dissertation, and finally I will explain the role that these methods play in each of the case studies.

### Fields of Study

Undertaking research at the crossroads of a technological and cultural domain, more specifically into data, race-ethnicity, and technologically mediated racialization, requires an interdisciplinary approach. That is the only way to make visible the specificities of both data infrastructures and race-ethnicity in the context of the Netherlands, and, more importantly, how they relate to and intersect with each other. In this paragraph, I will describe the scholarly fields of critical data studies and postcolonial studies, discuss how their combination will aid in my investigation.

### Critical Data Studies

As I already explained in the first chapter, processes of quantification, datafication, and computation are never neutral. The loose collection of studies that deals with this non-neutrality, and with the implicit and explicit political characteristics of data and data systems, has come to be known as *critical data studies* (CDS). This field of studies should be seen as

a formal attempt at naming the types of research that interrogate all forms of potentially depoliticized data science and to track the ways in which data are generated, curated, and

how they permeate and exert power on all manner of forms of life. (Iliadis and Russo 2016, 2)

I want to explicitly place myself in this relatively young research tradition of CDS, by showing the forms of power exerted by governmental data systems. Moreover, my research methodologies consisting of discourse analysis, affordance analysis and infrastructural inversion are firmly embedded within this field.

Neither technologically mediated racialization nor datafication happen in a vacuum; they are interrelated and contextually situated phenomena. An important strategy in the investigation of data systems, then, is to perceive of them not as individual machines or single technological objects, but as *assemblages* embedded within a particular context (Dalton and Thatcher 2014; Dalton, Taylor, and Thatcher 2016; Iliadis and Russo 2016). A *data assemblage* bears similarities to a Foucauldian *dispositif*, as discussed in Chapter 1, in that it refers to a particular system of knowledge production and dissemination, embedded in its physical, social, institutional, legal, and philosophical context. *Data assemblages*, too, are thought of as collections of elements which may include

systems of thought, forms of knowledge, finance, political economy, governmentalities and legalities, materialities and infrastructures, practices, organizations and institutions, subjectivities and communities, places, and the marketplace where data are constituted. (Kitchin and Lauriault 2018, 3–4)

Data assemblages, therefore, are subjected to vast socio-technical systems, which are grounded in “engineering and industrial practices, technological artefacts, political programs, and institutional ideologies which act together to govern technological development” (Kitchin and Lauriault 2018, 19). It is this notion of data assemblages as systems of thought, forms of knowledge, and infrastructures that can provide us with a more historically and discursively aware understanding of how power and hegemony work through data technologies such as the ones under scrutiny in this dissertation.

My investigation of infrastructures as data assemblages will translate into analyses of categories, definitions, labels, numerical values associated with categories, interfaces, and visualizations. Furthermore, I will situate these technical systems in their larger organizational and cultural context, as well as in the larger material infrastructures within which they are embedded. Practically, this means that for all data under scrutiny in this dissertation, I will look both at how data are distributed by means of an infrastructure, and how this data is appropriated for a specific purpose in the applications in which they end up being used. Practically, this means that I take into account *cultural* notions—what counts as a “migration background”—and *material/technological* notions—that is, how nationally sanctioned categories such as “person with a migration background”, or “allochtoon”, travel from and infrastructure to a specific digital

application. The purpose of this is to understand how racializing meanings of categories are created, attributed, changed, and lost in the process of datafication. Through the integration of these analyses, I will be able to show that technologically mediated racialization in the context of the Netherlands is not incidental, or located only in specific data applications, but rather a process that happens throughout the infrastructure, its attached applications, and the context in which these technological systems do their work.

Through a critical data studies perspective on technologically mediated racialization, several different, more general, types of critique can also be formulated (Dalton and Thatcher 2014; as listed in Kitchin and Lauriault 2018, 15):

- Situating data regimes in time and space;
- Exposing data as inherently political and identifying whose interests they serve;
- Unpacking the complex, non-deterministic relationship between data and society;
- Illustrating the ways in which data are never “raw” (see also Gitelman 2013);
- Exposing the fallacies that data can speak for themselves and that big data will replace small data;
- Exploring how new data regimes can be used in socially progressive ways; and
- Examining how academia engages with new data regimes and the opportunities of such engagement.

In the context of this dissertation, I will not only use these critical lenses to question the perceived neutrality and objectivity of governmental data systems in general, but focus in particular on the kinds of embedded power relations that relate to the Dutch postcolonial situation. To be able to properly engage with these issues, then, we need one more critical lens: postcolonial studies.

## Postcolonial Studies

The field of postcolonial studies focuses on the ongoing influence of imperialism and colonialism on culture, philosophy, politics, and history (Ashcroft, Griffiths, and Tiffin 1989; 1995; McLeod 2010). In the words of the influential postcolonial theorist Robert J.C. Young:

At its simplest level, postcolonial theory results from different experiences of cultural and national origins, the ways in which the colour of the skin affects anyone’s life in the metropolis, the ways in which your place of birth determines the kind of life, privileged and pleasurable, or oppressed and exploited, you will have in this world. (Young 2004, 7)

It is important to note that, within the field of postcolonial studies, coloniality is not merely tied to a particular time period or spatial territory, like colonialism, but should be understood as a psychological, political, economic, social, epistemic, or ontological condition. Where colonialism describes political and economic power relations, coloniality should be seen as “long-standing



patterns of power that emerged as a result of colonialism, but that define culture, labor, intersubjective relations, and knowledge production well beyond the strict limits of colonial administrations” (Maldonado-Torres 2007, 243). Postcoloniality, then, is both an affirmation of the ongoing power of historical hierarchies and power relations, and an attempt to move beyond them (Hall 1995).

There has been some discussion on whether or not the word should include a hyphen between “post” and “colonial”. In *The Empire Writes Back*, literary scholars Bill Ashcroft, Gareth Griffiths and Helen Tiffin state about the hyphenated version that

[g]rounded in the practice of critics concerned with the writing of colonized peoples themselves it has come to stand for a theory oriented towards the historical and cultural experience of colonized peoples, a concern with textual production, rather than towards the fetishization of theory itself. (Ashcroft, Griffiths, and Tiffin 1989, 198)

In opposition to this notion, John McLeod (2010), notes that “post-colonial” seems more suited to signify a time period, an epoch, rather than an analytical framework. “Postcolonialism”—without the hyphen—denotes “something that one does: it can describe a way of thinking, a mode of perception, a line of inquiry, an aesthetic practice, a method of investigation”(McLeod 2010, 6). Additionally, Kwame Anthony Appiah notes that “the post- in postcolonial, like the post- in postmodern, is the post- of the space-clearing gesture”, meaning that is “concerned with transcending, with going beyond, coloniality” (Appiah 1991, 348). Since this research takes postcolonial critique into the study of data systems, I do not consider the unhyphenated word “postcolonialism” a form of theory fetishism, but rather as a lens through which epistemological and ontological consequences can be diagnosed and analyzed.

Moreover, a postcolonial studies perspective allows for a focus on race-ethnic power hierarchies that goes beyond traditional views on race, ethnicity and nationality in terms of skin color or other visible characteristics people might have. Thinking in terms of colonial relations

moves beyond the tendency in racial/ethnic studies to focus only on the persistence of a color hierarchy. (...) Moreover, a focus on color alone does not address the fact that, although diverse colonized groups may be phenotypically indistinguishable from dominant colonizer groups, they can nevertheless be racialized as inferior others in a colonial situation. (Grosfoguel 2004, 326)

Such a “colonial situation” does neither limit itself to people with origins in former colonies of a nation, nor does it simply refer to people with a darker skin tone. In the Dutch context, we should, therefore, not only consider people from former Dutch colonies, such as Suriname and Indonesia, “colonial immigrants”, but also so-called “guest workers” who moved from Morocco and Turkey in the 60s and 70s (see Grosfoguel 2004, 327). A postcolonial perspective makes the “invisible visible” (Young 2012, 23); it highlights patterns of dissonance between colonial legacies and the

postcolonial predicament. Therefore, it allows for a more inclusive analysis of how governmental bureaucratic infrastructures create and maintain hierarchies between people with different characteristics that relate to ancestry, whether it be in terms of skin colour, ethnicity, nationality, or any other characteristic that is used in practices of racial signification.

Within the context of this dissertation, a postcolonial perspective on datafication means that special attention is given to the power asymmetries stemming, albeit transformed, from colonial hierarchies. As I will show in my introduction to the case studies in the next paragraph, this means that I will look into the definitions used for different race-ethnic categories, *and* their related social and technical historical origins. In addition, through a postcolonial lens, I will be able to show how the politics of the affordances of data systems are creating and perpetuating a colonial situation in the Netherlands. When we place the aforementioned race-ethnic categories and the racializing affordances of individual technologies within the data infrastructures of the Dutch government it will, again, become clear that technologically mediated racialization is not incidental but rather an infrastructurally embedded and, therefore, institutional phenomenon.

## Interventions

Before I continue with a discussion of the main research methods of this investigation, as well as the introduction of the case studies, I will briefly discuss which interventions I am going to make in the aforementioned fields of study. It is my aim to not only add new directions and depth to these individual research fields, but also show how the *specific combination* of critical data studies and postcolonial studies can be made productive in the study of technologically mediated racialization. This dissertation, then, aims to provide three main interventions. I will 1) contribute to the ongoing implementation of more ontological perspectives on datafication in the field of critical data studies, a field that is still largely concerned with epistemological critique, 2) add an ontological understanding of technology in general, and data in particular, to the field of postcolonial studies, and 3) show how the combination of both critical data studies and postcolonial studies is essential in the study of technologically mediated racialization. In the next few paragraphs I will explain these interventions in more detail.

First, critical data studies has, from its inception, mainly focused on epistemological arguments in its critique of rather instrumentalist understandings of datafication in fields where it was first implemented (see boyd and Crawford 2012; Kitchin 2014a; Dalton, Taylor, and Thatcher 2016; Iliadis and Russo 2016). Such epistemological critique focusses on the ways in which data and algorithms represent, or misrepresent, a world that is considered to exist independently of the ways in which it is known. Obvious signifiers of epistemological arguments are words that remain stuck in a split between a world that is *known* through data and a world that is *conceived* through and with that data (i.e. “mediated” as understood through the ontological approach forwarded in Chapter 1), such as “data double” (see Haggerty and Ericson 2000), and “data derivative” (see

Amoore 2011). Each of these concepts points at the ways in which data about a person are separate from that person and the implications of that separation for surveillance practices. In the few investigations where ontological arguments are made—showing how, increasingly, people, things and phenomena, do not exist outside their technological mediation—, they often remain rather theoretical and abstract (see Iliadis 2013), or rather implicit (see Kitchin 2021). John Cheney-Lippold is one of the few scholars whose work makes a theoretical argument about the ontological implications of data technologies, and connects this in a productive manner with empirical case studies on social media and surveillance technologies of the U.S. government. With this dissertation, I aim to contribute to this ontological argument by showing how data technologies mediate processes of racialization in contexts of government bureaucracy in the Netherlands, and on a more generalized level, how processes of datafication are not merely representative, but also generative, of contemporary conceptions of race. Furthermore, by showing different types of approaches to this problem, it becomes possible to go beyond mere critique, and provide practical advice that can be used in the implementation of future data technologies.

Second, investigations within the field of postcolonial studies that deal with data and datafication mainly criticize how various stereotypes and racializing epistemologies persist in the digital domain (see Nakamura 2008; 2013; McPherson 2012; Sharma 2013; Daniels 2013; Browne 2015; Noble 2018; Eubanks 2018; Risam 2019; Benjamin 2019). In addition, the colonial attitude of companies that extract data and sell their services is addressed in business and media literature (Zuboff 2015; Couldry and Mejias 2019a; Madianou 2019). Furthermore, the field of postcolonial science and technology studies tends to focus on the persisting imperial and colonial influence in various sciences (see W. Anderson 2002; Adams 2002; McNeil 2005; Abraham 2006). With this dissertation, I aim to contribute to the investigation of power structures present in data technologies and show how it can benefit greatly from postcolonial critique. Through the case studies of Chapters 3, 4, and 5, I will show that postcolonial studies and critical data studies are mutually informing each other, co-constituting the critical field of postcolonial critical data studies. I will show that this combination is rather essential in understanding the scope and inner workings of technologically mediated racialization through data technologies in the Netherlands.

Third, while both critical data studies and postcolonial studies are largely aligned in their aims to expose power structures embedded in knowledge producing systems, projects that combine the two perspectives in the study of technology remain relatively scarce. Critical Data Studies that employ postcolonial and decolonial critiques mainly criticize the rather Western bias of critical data studies (Arora 2016; 2019; Milan and Treré 2019; Segura and Waisbord 2019; Ricourte 2019; Gangadharan and Niklas 2019). Other critical data studies with a postcolonial sensitivity can be found in the field of digital migration studies where they mainly focus on the digital interactions between refugees, and between refugees and the countries that monitor them (Dijstelbloem and

Meijer 2011; van der Ploeg and Sprenkels 2011; Fassin 2011; Leurs and Shepherd 2017; Leurs and Ponzanesi 2018; Leurs and Smets 2018; Sánchez-Querubín and Rogers 2018; Madianou 2019; Ponzanesi 2019; Dijstelbloem 2021; Al Jaramani, Ponzanesi, and van Schie 2022). Much of this research deals with the role of various social media migrants use in telling their stories, and/or various surveillance technologies employed by countries to discourage and prevent migrants from entering their borders. Postcolonial critical data studies about data technologies employed to scrutinize people that are already legal citizens of a particular country remains relatively scarce (for a rare exception, see Isin and Ruppert 2019), and are for the specific case of the Netherlands, still missing. Therefore, my main focus in this dissertation is the persisting colonial situation in the production of Dutch and non-Dutch identities and how these are mediated, constituted and raced by the network of infrastructures and data technologies in the context of the Dutch government.

## Methods

My investigation into the instrumental, epistemological, and ontological implications of technologically mediated racialization in Dutch governmental data systems is informed by both critical data studies and postcolonial studies. I am, therefore, interested both in technologies, their cultural context, and the relationship between the two. As a result, my analyses are not limited to either objects, descriptions, or networked connections, but include both language, technologies, and infrastructures. It is for this reason that I employ three different research methods in conjunction with each other: critical discourse analysis, affordance analysis, and infrastructural inversion. While the technologies and contexts under scrutiny will differ, I will consistently start with a critical discussion of the race-ethnic categories used in a particular data system. Through critical discourse analysis, I will look at the implicit and explicit meanings of categories, as well as the embedded power relations that emerge through their organizational and cultural history. With an affordance analysis, I will be able to study the implementation and functionalities of race-ethnic categories in particular data technologies. Practically this means that I will investigate how hardware and software enable users of systems to work with race-ethnic categories, and, as a result, provide possibilities for technologically mediated racialization. Finally, through an infrastructural inversion, it will be possible to show how different data systems relate to each other, creating possibilities for race-ethnic categorization and racialization to travel from one context to another. Throughout the analyses of technologically mediated racialization in the coming chapters, I will use a mixed method approach combining the methods discussed below. Since meaning, functionality, and material infrastructures are related to such a degree, sequencing individual methods would not make much sense and would also create too much repetition for the reader.

## Critical Discourse Analysis

Critical discourse analysis (CDA) deals with the investigation of discourse, which, in the context of this dissertation, is understood as “a broad conglomeration of linguistic and non-linguistic social practices and ideological assumptions that together construct or reinforce power” (Tannen, Hamilton, and Schiffrin 2015, 1). Where cultural studies research on discourses mainly focused on media's representations of social issues (see for example Hall et al. 1978) and was very often preoccupied with textuality<sup>13</sup> (see Hall 1992, 271), CDA also explicitly includes institutional and sociocultural contexts (Carvalho 2008). It should be noted that CDA is not technically a method, but rather an analytical practice, that might incorporate most methods available to social scientists and humanities scholars (see van Dijk 2015, 466). In this dissertation, I will enhance this analytical practice the previously discussed perspectives of critical data studies and postcolonial studies. Since the power relations that CDA aims to reveal and explain generally do not merely reside in spoken or written text, CDA should *not* be seen as “analysis of discourse ‘in itself’ (...), but analysis of dialectical *relations between* discourse and other objects, elements or moments, as well as analysis of the ‘internal relations’ of discourse” (Fairclough 2013, 4). Within this dissertation, I am therefore not only interested in what is communicated in data systems, but also in policy documents, evaluation documents, census reports, and websites that relate to the systems under scrutiny. In addition, I will not only be interested in text, but also in pictures, graphs, tables, maps, interfaces, and other more visual ways in which data, computation, and results are communicated with researchers, citizens, journalists, and other interested users of systems.

Reading and describing the available information on the use of race-ethnic categories in data systems and infrastructures is usually enough to properly conduct an investigation within the instrumental approach to technologically mediated racialization. In each of the cases, I will read the research and evaluation reports, looking into specific reasons for the use of race-ethnic categories. If available, I will investigate why race-ethnic categories are implemented in a data technology, and which values are associated with these categories. Therefore, I will examine the phenomena they are cross-referenced with, such as crime or welfare fraud, and what values are attached to a category in code. However, merely taking available information at face value does not constitute a *critical* discourse analysis. For the epistemological approach, we need to also address power relations and normative assumptions embedded in language and policy choices. The starting point for each of the analyses will be the race-ethnic categories that are present in a system. Each time, the aim will be to achieve a thorough understanding of both the implicit and

---

<sup>13</sup> In cultural studies, media studies, and several related fields, “textuality” and “text” can refer to any mediated message including sound, imagery, video, etc.

explicit meanings of categories such as “allochtoon”, “autochtoon”, “foreigner”, “person with a migration background”, and of how these categories relate to the specific system in which they are used. In my historical analysis in Chapter 3, I will use census reports and accompanying documentation to find out which categories are used and how they are clustered and defined. When these systems are web based applications, like the Allochtoon-o-meter and Leefbaarometer in Chapter 4, I will look for definitions and explanations about the used categories provided on websites in which these systems are embedded. Additionally, two of the systems under investigation in this dissertation, the Leefbaarometer, and the case study of Chapter 5, the Crime Anticipation System of the Dutch National Police, are well-documented government systems which were tested and evaluated multiple times. By reading the available information through both an instrumental *and* an epistemological approach to technologically mediated racialization, we will be able to distinguish between the official and intended policy choices and their ideological and normative implications.

An additional way of showing the difference between the instrumental and epistemological approach in the investigation of data systems is to show that many systems have an embedded norm concerning datafied information about citizens. To assess how data systems imply such a norm we can look at situations in which the culturally dominant category is not represented at all. Donna Haraway’s (1991, 188) remarks about the unmarked category claiming “the power to see and not be seen”, and “to represent while escaping representation”, prove to be particularly informative in this regard. As I will show in each of the web based applications and the Crime Anticipation System, “autochtoon” or “person with a Dutch migration background”, are treated as default reference categories, about which the makers do not feel the necessity to include them in tables or code. Such instances, in which the culturally dominant category is not explicitly made a part of knowledge production or knowledge dissemination, are not limited to data systems, but are rather a more general feature of discourses on migration and integration (see Hartigan 1997; Emirbayer and Desmond 2012; Mepschen 2016, 29). In making these embedded values explicit, we can show how the data systems under scrutiny are not isolated instances of technologically mediated racialization. Their processes are connected with broader cultural values and attitudes concerning race that are present in Dutch society. Through discursive analyses of categories, interfaces, and other documentation concerning specific data technologies, we can show how the rather positivist and empiricist assumptions of an instrumental approach to technologically mediated racialization lack cultural and political context. The approach to categories as objective descriptors, and data technologies as neutral instruments lacks an awareness of how hegemony and power work their way into the production of knowledge about groups of people. The tools that aid in knowledge production should be considered active agents in this process. Through discursive analyses, we can take into account both explicit and implicit meanings attached to categories, and the way in which they aid in the production and maintenance of power relations through technological systems created to produce knowledge about a population. When used in

this his fashion, CDA can be considered an important tool in the epistemological approach to technologically mediated racialization.

While CDA aids in the investigation and explication of discourses on race-ethnicity in technology we need additional methodological tools in order to 1) situate discourses within their “discourse networks”, 2) connect particular racial formations to the particular functionalities of the technological objects with which they are entangled, and 3) place these discourses and technological objects within larger infrastructures. Apart from merely representing people, categories in statistics “make up” people (Hacking 1986). Categories do not only have a discursive dimension, but also a performative dimension; categories *mean*, as well as *create*. A next step in the analyses of racialization in data systems, then, would be to include not only implicit and explicit meanings of categories in text, but also the structures and functionalities of data technologies. Therefore, for the ontological approach to technologically mediated racialization we need to know more about the functionalities of data technologies and their embeddedness in larger infrastructures. Such an inclusion asks for an upgrade of a rather Foucauldian discourse analysis into the analysis of what media scholar Friedrich Kittler (1990) calls *aufschreibesysteme*, or in English “discourse networks”, which he defines as “the network of technologies and institutions that allow a given culture to select, store, and process relevant data” (Kittler 1990, 369). While Kittler’s *aufschreibesysteme* mostly describe technological infrastructures that predate the implementation of large scale digital information infrastructures such as, for example, television, it still allows us a perspective on data that goes beyond the individual data technologies and makes visible how they operate in infrastructures. It is this conceptualization of discourse networks that will allow me to not only look at racialization as the content of data technologies, but also understand racialization as a process integral to the functioning and operation of infrastructures. The selection of human characteristics for racial signification, then, is not only related to culture, but also to the *technical possibilities* technologies and infrastructures have to offer. Therefore, before we move to a discussion of how infrastructures can be investigated, we first need to understand how such technical possibilities can be made the subject of an investigation through the concept of affordances.

### **Affordance Analysis**

The functionalities of technological objects have been theorized by media and design scholars in terms of affordances (see Norman 1999; Hutchby 2001; Chemero 2003). Affordance theory offers a way into the analysis of systems and parts of systems as functional entities. The concept *affordance* originates in the work of ecological psychologist James J. Gibson, who used the word *affordance* to talk about the possibilities of an environment for the animals that inhabit it. In this sense he explicitly understood affordances in relational terms:

The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson 1979, 127)

While Gibson understood affordances as a relationship between an animal and its environment, in design theory this term was appropriated to more exclusively designate the possibilities an object has to offer a user. In his book *The Design of Everyday Things*, the cognitive psychologist and design theorist Donald A. Norman (1988) describes affordances as

the relationship between a physical object and a person (or for that matter, any interacting agent, whether animal or human, or even machines and robots). An affordance is a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used. (Norman 1988, 11)

Technological objects are selected to be used in bureaucratic practices for their specific functionalities. Each of these functionalities also makes other uses harder or even impossible. A proper analysis, therefore, takes into account both the possibilities and impossibilities caused by particular designs, materials used, and the ways in which technologies are embedded in existing infrastructures.

The affordances of technological objects that I will be analyzing in the case studies of this dissertation focus on the material and immaterial aspects of the technologies in question. I will look at these affordances at three different levels, namely during collection, dissemination, and appropriation. On the level of data collection, I will question how race and ethnicity are operationalized, and used for computation. Practically, this means that I will investigate which human characteristics are used as the basis for race-ethnic categorization and how these characteristics relate to the design characteristics of the counting technologies used in the censuses, administrations, and databases of Dutch municipalities. On the level of data dissemination, I will look specifically at how different race-ethnicities are clustered and made available for dissemination by Statistics Netherlands. The nationality based system of the Netherlands does not produce statistics about almost 200 existing nationalities, but rather uses terms like “autochtoon”, “allochtoon”, “person with a migration background”, “Western”, and non-“Western” to cluster different countries, after which the aggregated information is made available through an API (*application programming interface*). These aggregated data can then be appropriated by other systems to process even further. The selected systems under scrutiny in this dissertation all use data made available by CBS. In these systems, I will look at what design choices with regard to race-ethnicity are already present in CBS data, and what design choices were specifically made in systems like the Allochtonenmeter, Leefbaarometer, and Crime Anticipation System. The types of questions I will attempt to answer are:



- Which technological affordances make race-ethnicity countable and how?
- Which technological affordances make race-ethnicity available for dissemination and how?
- What affordances of technologies are directly related to the appropriation of race-ethnicity, or in other words, what can systems do *because* of their use of race-ethnic characteristics?

Taken together, these questions can be used to map not so much the *meanings* of race-ethnicity in systems, but rather the way in which they *perform*. It is this operationalization of race-ethnicity that will provide more insight into the ontological dimension of technologically mediated racialization. In addition, it shows how the availability of race-ethnicity itself provides particular affordances to technologies built on an already available infrastructure. To achieve more insight into the infrastructural nature of technologically mediated racialization in the Netherlands we, therefore, need one other method: infrastructural inversion.

### Infrastructural Inversion

Apart from studying the discourses and affordances entwined with the datafication of race-ethnicity, we have to examine the infrastructures in which they are embedded. As large systems governed by rules and standards, infrastructures are made in such a way that working with them ideally happens effortlessly and without overthinking things. Investigating these infrastructures can prove to be rather difficult due to their tendency to remain invisible to us. They sometimes quite literally “fade into the woodwork” (Bowker and Star 1999, 34). Try to think about the moment you last charged your smartphone. Did the power plug fit in the socket? Did the voltage provided by the energy company match the one your phone charger needed to function properly? The reason that all these things work without you having to check every single time is due to the infrastructural nature of the power grid. To get to this point, many standards had to be negotiated and set to make sure that all governments and companies in a given area work with the same standards. Standards can be very material, such as the exact dimensions of a power plug or socket. Or rather immaterial, such as the voltage provided by the energy company or the informational standards that allow your laptop to communicate with the WIFI hotspot. The specific format chosen for each of these standards is necessarily a compromise, not only between functionality, available materials and cost, but also between sociocultural, political, and economic forces.

One of those important sociocultural forces is that old technologies pave the way for newer ones, a process which is called path dependency. Path dependencies emerge when subsequent technologies adhere to a form that was previously established, either due to social and technological habituation, switching costs being too high, or both (Berkhout 2002). A famous example of such a path dependency is the use of the QWERTY-keyboard, which was created in

the second half of the 19<sup>th</sup> century as the best layout to prevent the physical letters of a mechanical type writer from clashing into each other when typing sentences in English (David 1985). During the past century, 1) other layouts turned out to be much more efficient for human hands (see David 1985, 332), 2) keyboards started to no longer feature mechanical components capable of clashing, and 3) keyboards started to be used for many other languages apart from English. These are all very good reasons to switch to other systems, but the QWERTY layout remains the standard. This example shows that the choice of technological standards is not necessarily guided by what is the best or most efficient, but rather takes into account what other actors in an infrastructure, that is, both people and machines, are already capable of working with.

Similar to phone chargers and keyboards, the production of knowledge is guided by many infrastructures and standards, each with their own types of path dependencies and practical politics. Mice are chosen over elephants as lab animals not for epistemic reasons, but rather for their low cost and easy handling. Furthermore, some countries in the world still choose the English imperial system of measuring units rather than the metric system, even though this causes international misunderstandings and expensive mistakes (see Lewis 2015; Enderson and Selover 2021). In a similar fashion, the choices of categories in Dutch governmental data systems are not only made on the basis of epistemic rigour, but also simply on the basis of what is already available, what is cheap, and what people are used to. The final methodological tool that I will employ in this dissertation - infrastructural inversion - focusses exactly on those practical politics of the construction of an infrastructure and its associated standards. An infrastructural inversion is explained by STS scholar Geoffrey C. Bowker and sociologist Susan Leigh Star as recognizing “the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other” (Bowker and Star 1999, 34). This methodology, invented for and used in research on knowledge production in science and policy, starts from the very basis on which a system is built: the categories and definitions that are the foundation of an infrastructure and the historic, cultural, and material contexts in which they emerge.

In the context of this dissertation, an infrastructural inversion is seen as a method to trace the mutual shaping of categories and epistemic technologies and practices in technological infrastructures, both through time and through the different connections between technological systems within a network. An infrastructural inversion is, therefore, a methodological process that enables us to critically engage with technical and informational standards in systems and, more importantly, examine how the current informational standards were negotiated over time. By tracing categories back from the moment we encounter them, to the moment they were assigned—this would be the *inversion*—I will be able to show how technologically mediated racialization does not happen within a singular system or at a particular place, but rather happens

throughout the Dutch governmental data infrastructure. Practically, this means that I will begin by examining the technological objects users encounter, such as the Allochtonenmeter, Leefbaarometer, and Crime Anticipation System, after which I will systematically trace their data sources, as well as the categories and definitions used in those sources. After this step, I will be able to trace the data one more time, as CBS—the main data source for these projects—does not collect the data themselves, but rather taps into the bureaucratic systems of Dutch municipalities. Here, again, I will look into the informational standards and ways in which race-ethnicity is assessed. When taken as a whole, the product of an infrastructural inversion is a map of the different paths data take from the moment of collection, through transport, all the way to the moment of dissemination. Included in this map, are the different shapes and forms race-ethnic information takes in terms of categories, labels, definitions and clusters, depending on the affordances of the technologies in which they exist.

Taken together, critical discourse analysis, affordance analysis, and infrastructural inversion provide a solid basis to not only critique instrumentalism, but also go beyond a merely epistemological understanding of technologically mediated racialization. As entwined phenomena, data technologies and race-ethnic categories, create a particular technicity of race in the context of the Netherlands. Revealing this technicity of race will aid us in understanding how the data technologies and the race-ethnic categories used in the 70s and 80s have led to the current situation in which a centralized database of Statistics Netherlands functions as the infrastructure for other technologies employing various forms of technologically mediated racialization. In addition, seeing the current situation in terms of technicity can be very informative in guiding future policy concerning the usage of race-ethnic categories in data technologies, as well as informing design choices to prevent technologically mediated racialization to become a part of the bureaucratic systems and infrastructures of tomorrow. In the next paragraph, I will introduce the case studies, focusing not so much on instrumental, epistemological, and ontological issues with bureaucratic systems in general, but rather on the particular situations that arise when data technologies are combined with race-ethnic categories and racializing affordances in a postcolonial context.

## Conclusion

In the next three chapters, I apply the insights from the fields of critical data studies and postcolonial studies and the methods introduced in this chapter. The topics of these chapters are in order of appearance:

- Chapter 3: A combined history of race-ethnic categorization and counting, calculation and dissemination technologies from 1899 to 2021;

- Chapter 4: A comparative analysis of two web-based applications, the Allochtonenmeter and the Leefbaarometer, both based on data that is made available by the infrastructure provided by CBS;
- Chapter 5: An analysis of the Crime Anticipation System, a predictive policing system developed and implemented by the Dutch National Police.

I have selected each of the case studies central to these chapters for different reasons. In the coming paragraphs, I will introduce each of the cases and explain the role of the different methods, as well as their relation to the instrumental, epistemological, and ontological approaches to technologically mediated racialization outlined in the first chapter of this dissertation.

Each of the three chapters with case studies has a different role in the larger picture of this dissertation, and the three main methods produce different results in each of them (see Figure 4). Where critical discourse analysis in Chapter 3 aids in linking different categorization practices and labels with national social policy from the start of CBS until now, in Chapter 4 it aids in showing how race-ethnic categorization has become naturalized in questions concerning demography. As a result, it does not only show up in a system that is part of alt-right rhetoric, but also in a system presented on an official website of the Ministry of the Interior of the Dutch government. In Chapter 5, I will employ critical discourse analysis in reading the CAS evaluation reports and show how they imply rather technologically deterministic visions of the role that technology will play in future policing.

In Chapter 3, affordance analysis will aid in the analysis of counting and calculation practices that have been informing the different ways in which race-ethnic categorization was practiced in the Netherlands between 1899 and 2021. The development of computer technologies and the possibilities of connecting different databases from the 1980s onwards created the possibility for third parties to use data for different purposes than merely monitoring the effects of social policy. In Chapter 4, the combination of discourse analysis and affordance analysis will show how different systems communicate or obfuscate the perspective from which their knowledge is produced. This allows us to reflect on the different ways in which systems such as the Allochtonenmeter and Leefbaarometer imbue particular race-ethnic categories with new and more intricate racializing meanings. In Chapter 5, affordance analysis will be used in a similar fashion with the added aim of connecting affordances with police culture. Through the concept of the data assemblage, the results of the affordance analysis will be combined with recent ethnographic and sociological research on the organizational culture of the police in general and ethnic profiling in particular. This has the added benefit of understanding data systems as situated knowledge producing practices that are both intricately linked to the particular

organization in which they are employed, as well as to the larger social, cultural and technological context in which they are embedded.

Together, these case studies are part of a larger infrastructural inversion that traces back the roots of race-ethnic information from the moment that a user encounters an interface that shows the results of a data system or online application. By systematically tracing back how data are visualized, used for calculation, and collected, I am able to retrieve connections that have become lost or obfuscated through label changes and technological interfaces. In this way I aim to show how contemporary practices of datafication that relate to the migration backgrounds of Dutch people are grounded in a longer history of racializing bureaucratic practice. For clarity and readability, I will present this infrastructural inversion in chronological order, starting in the past when CBS was founded, and working forward to the present in which CBS has become the central node in the governmental data infrastructure that supports a wide range of data application making use of race-ethnic information.

	<b>Critical Discourse Analysis</b>	<b>Affordance Analysis</b>	<b>Infrastructural Inversion</b>
<i>Chapter 3</i>	Mapping the different words that have been used to signify Dutchness and Otherness in the past century and linking them to historical developments.	Connecting the development of the affordances of technologies to the changing definition and operationalization of race-ethnic categories.	Mapping the history and emergence of contemporary race-ethnic categories and associated technologies as a starting point for the case studies of Chapters 4 and 5.
<i>Chapter 4</i>	Linking the contemporary use of race-ethnic categories and associated technologies to political and historical developments in the first decade of the 21 <sup>st</sup> century.	Showing the ways in which affordances affect the situatedness of knowledge production in the datafication of race-ethnicity.	Connecting the race-ethnic categories in the Leefbaarometer and Allochtonenmeter to the CBS infrastructure and by extension to the history discussed in Chapter 3.
<i>Chapter 5</i>	Linking the discourses on race-ethnicity and datafication with the specific national and organizational culture of the Dutch police. Special attention is given to the implications of the combination of ethnic profiling practices and a rather technologically deterministic vision of the Dutch police	Situating the affordances of CAS within the culture of police organization. Including reflections on absent information and possibilities, which reflect a disinterest in processes of accountability concerning the equal treatment of Dutch citizens.	Connecting the race-ethnic categories in CAS to the CBS infrastructure and by extension to the history discussed in Chapter 3.

Figure 4: Function of the three main methods in each case study



## Chapter 3: The Origins of Race-Ethnic Categorization in the Dutch Governmental Data Ontology (1899-2020)

Origin (n.): c.1400, ‘ancestry, race,’ from Latin *Originem* (nominative *origo*) ‘a rise, commencement, beginning, source; descent, lineage, birth,’ from stem of *oriri* ‘arise, rise, get up; appear above the horizon, become visible; be born, be descended, receive life.’<sup>14</sup>

The race-ethnic categories at the basis of many economic and social statistics produced by *Centraal Bureau voor de Statistiek* (Statistics Netherlands, from now on CBS)<sup>15</sup>, as well as subsequent data systems built in relation to this data infrastructure, are the result of an ongoing process of historical, social, and technological developments. In order to better understand the discourses and discourse networks of the race-ethnic categories currently used in the Netherlands, this chapter will function as an infrastructural inversion (Bowker and Star 1999, 34) of the common CBS categories “*autochtoon*”, “*allochtoon*”, “*Westers*”, and “*niet-Westers*”, used roughly between 1970 and 2021. By looking closely at “technologies and arrangements that, by design and by habit, tend to fade into the woodwork” (Bowker and Star 1999, 34), I will discuss both the meanings and the functionalities of the race-ethnic categories used by CBS in the past century and in doing so I will go beyond the time period in which the “*autochtoon*”/“*allochtoon*” dichotomy was the official vernacular of the Dutch government. I will show the continuities and discontinuities in race-ethnic classification from 1899, the year in which CBS was founded, until the period 2017-2021, when this investigation took place, when the Dutch government abolished the aforementioned categories and began to use the word “migration background” to designate race-ethnic differences. Through my focus on the technological aspects of categorizations, I will show how, in line with technological advances in other areas of life, processes of racialization receive regular updates (Betancur and Herring 2013; Browne 2015; Benjamin 2016; Dixon-Román 2016; 2017; Saini 2020). While such updates are usually intended to improve bureaucratic accuracy in relation to institutional and policy needs, they simultaneously cut off racializing discourses from their roots and mask continuities in racializing practice. In this chapter, I aim to undo such technologically mediated “colonial aphasia”, which is understood as the “dismembering of words from the objects to which they refer, a difficulty retrieving both the semantic and lexical components of vocabularies, a loss of access that may verge on active

---

<sup>14</sup> From the Online Etymology Dictionary: <https://www.etymonline.com/word/origin>.

<sup>15</sup> The literal translation of *Centraal Bureau voor de Statistiek* is Central Bureau for Statistics. Its acronym CBS is well-known in the Netherlands. By sticking to the Dutch acronym I do not only save space, but also make the text easier to read for a Dutch audience.



dissociation” (Stoler 2016, 12).<sup>16</sup> Furthermore, in relation to the quantification and datafication of race-ethnicity in the Netherlands, I will show how different methods of classification have evolved 1) together with the technologies and infrastructures available, 2) in relation to institutional and policy contexts, and 3) in relation to the larger geopolitical and societal context of the discourse on race-ethnicity in the Netherlands. This way, I will be able to address the first two subquestions of this dissertation: This way I will be able to answer the first two subquestions of this dissertation:

1. What are the origins of the Dutch race-ethnic categorization system and how are its categories defined?;
2. How are race-ethnic categorization systems in the Netherlands datafied by government institutions?

The specific set of practices and technologies analyzed in this chapter, were used during census taking in the Netherlands between 1899 and 2011. As sources, I use the census reports produced by CBS between 1899 and 2020—that is, those of 1899, 1909, 1920, 1930, 1947, 1960, 1971, 2001, and 2011.<sup>17</sup> In his influential book *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, Benedict Anderson (2006) describes the census, together with the map and the museum, as a particularly important medium in the construction of the modern nation-state and its connection to race, ethnicity, and religion. Apart from its role in the formation and development of European nation states (see Emigh, Riley, and Ahmed 2016a; 2016b), census taking was also instrumental in the discursive construction of the colonial Other through registering racial and ethnic categories (B. Anderson 2006, 164–70). Furthermore, in the context of this study, it should be noted that the practice of taking the census is rather intimately connected with the social and technological infrastructures available, since

the available forms of social organization might determine whether you could count at all. A complete census of a large population requires sophisticated bureaucratic structures, which few states possessed before the nineteenth century. (Porter 1995, 35)

---

<sup>16</sup> See the more detailed discussion on “colonial aphasia” in Chapter 1.

<sup>17</sup> Censuses were conducted once every ten years, with the exception of 1941 (due to World War II), 1981 and 1991 (both due to the fact that census taking was no longer necessary; see the remainder of this chapter). The census was reinstated in 2001 because the European Union expected counts from its member states. Since reports were produced every decade, they also provide insights into their relation to key social, technical, and sociotechnical changes over time.

Thus, since census taking is determined by the available forms of social organization and its bureaucratic technologies, it necessarily follows the dominant ideological and institutional ideas of a particular time and place. Census taking is, therefore, an important intersection of the social and the technological, the political and the infrastructural, race and technology, and, as a result, an ideal starting point for the investigation of technologically mediated racialization.

In this chapter, I will focus specifically on 1) the parts of Dutch censuses that relate to origins of people in the form of race, ethnicity, nationality, and place of birth, 2) the ways in which their definition has been informed by ethnic, racial, and religious discourses in the Netherlands over time, and 3) the technical and institutional systems that are used for counting and calculation. In addition, where available, I will discuss policy documents and influential academic works that have shaped the Dutch discourse on race-ethnicity, migration, and integration. I will consistently understand the categorizations in terms of “race-ethnicity”, since this combinatory term best captures how ideas about race, ethnicity, nationality, and religion are deeply entwined (see also Yanow, Van der Haar, and Völke 2016). For example, while CBS currently uses the term ‘ethnic origins’, it always counts people as former citizens of nations, not as members of an ethnic group. This becomes painfully clear in cases of territorial disputes, such as, for example, with Kurdish Turks and Palestinians, who are counted as Turks and Israelis, respectively.<sup>18</sup> Through these practices, CBS contributes to policy discourses on otherness and foreignness, but also on more nationalistic discourses concerning nativeness and Dutchness as important elements of the Dutch nation-state. In this chapter, the Dutch census functions as a snapshot of the continuities and discontinuities of the evolution of dominant frames on Dutchness, migration, and accompanying modes of technologically mediated racialization in each decade.

By analyzing technologically mediated racialization in the census through an instrumental, epistemological, and ontological approach, I aim to bring forward the different levels at which racializing practices and technologies of quantification and statistics produce meaning and effects. As I will show throughout this dissertation, race-ethnic categorization seems to be an important part of the production of demographic data, given its usage in a wide range of policy situations, and in various different third party data applications such as the ones that will be discussed in Chapters 4 and 5. However, neither the categories used by CBS, nor the technologies that mediate their definition, clustering, counting, and dissemination have remained the same over the past 120 years. Through an instrumental approach, I will focus on how categories and concepts such as race, ethnicity, religion, and migration are defined and operationalized. Furthermore, such an approach allows me to reflect on how the chosen definitions and

---

<sup>18</sup> See the relevant table at CBS (2021c); when Israel is selected as a variable, it contains the combined numbers of people residing in both Israeli and Palestinian Territories.

categories align with experienced societal problems and envisioned policy goals. In this way, I can determine how chosen practices of quantification function within their own historical, societal, and policy context.

Through an epistemological approach, I will show that contemporary race-ethnic categories and data technologies do not reflect some kind of natural or inevitable way of performing statistics. Rather, they are local (Loukissas 2019) and situated (Dalton and Thatcher 2014; Bezuidenhout et al. 2017; Kitchin and Lauriault 2018) practices of “making up people” (Hacking 1986). These practices are intricately linked to mediating history, culture, and technologies. Through a history of race-ethnic categorization and its relation to political, cultural, social, and technological development, I will not only show how the current systems, technologies, and ways of thinking evolved, but also how the relationships between categorization practices and counting technologies evolved. This makes it possible to argue against the idea that statistical and policy practice only developed in a positive—i.e. becoming faster, more specific, and ultimately better—direction, but rather that they evolved laterally—i.e. becoming better at certain things, and worse at others, ultimately doing different things. In other words, an epistemological approach will enable us to assess qualitatively, how technologically mediated racialization has evolved through different technological and epistemic practices, and to recognize positive qualities in past practices and negative qualities in contemporary knowledge producing practices.

Through an ontological approach, I will show that race-ethnic categorization and its mediation are both entangled, as well as instable knowledge producing processes that change over time. I will argue that categorization systems should not only be thought of as epistemological tools, but as actors that actively shape reality (Bowker and Star 1999). In this capacity, race-ethnic categories, such as the recently abolished “allochtoon” and “autochtoon”, function as a ‘régimes of truth’ (Foucault 1980), that are not only enacted through governmental policy but also through scientific, political, and public discourse (Rath 1991; Prins 2000). Following sociologist and statistician Alain Desrosières, Rottenburg and Merry (2015, 15) explain how the circulation of categories within particular academic and bureaucratic domains causes them to “hold” and, usually temporarily, be accepted as real:

When the actors can rely on objects thus constructed, and these objects resist the tests intended to destroy them, aggregates do exist – at least during the period and in the domain in which these practices and tests succeed. (Desrosières 1998, 101)

In other words, if these aggregates work in an institutionalized practice, they will be accepted as real; it is their institutionalization that makes categories more than mere representations. They become actors that shape and define us as people as well as the world we live in (Rottenburg and Merry 2015). I, therefore, consider datafied race-ethnic categorization systems to be ontologies, rather than mere epistemologies. This results in an understanding of data as sociotechnical

constructions of reality, “defined in the context of a specific epistemic culture as sets of norms, symbols, human interactions, and processes that collectively facilitate the transformation of data into knowledge” (Kuiler 2014). As Desrosières (1998, 101) argues, categories and their reality do not “resist the tests intended to destroy them” forever. At some point, categories do no longer fit policy goals, context, or technological infrastructure and are replaced by new categories. By taking a historical approach to the investigation of the social, cultural, and technological conditions that shape race-ethnic ontologies, we will be able to better understand how race-ethnic categories emerge and are made real for the time they exist, and how different technologies contribute to their demise and discontinuation. In this chapter, then, we will look at how race-ethnic ontologies emerged, evolved and disappeared again in the context of data collection and dissemination by CBS over the past 120 years.

The categorization practices of the Dutch government have been discussed by scholars from a variety of different disciplines, but usually with divergent scopes or aims: as part of a particular policy context (Yanow and Van der Haar 2013; Yanow, Van der Haar, and Völke 2016), in terms of racial and ethnic connotations (see for example Essed and Hoving 2014a; Weiner 2014b; Wekker 2016) and in terms of their performance of inclusion and exclusion within Dutch society (see for example Schinkel 2013; Boersma and Schinkel 2015; Weiner 2015). In my own intervention, I would also like to place the ontology of the governmental information apparatus in its historical context. By discussing technological and societal developments together, I will make an argument for shifting the understanding of race-ethnic categorization in terms of the conceptual pair of “allochtoon” and “autochtoon” from social constructs to *sociotechnical* constructs. These constructs, as epistemological tools, also have performative power (Espeland and Stevens 2008): they are capable of shaping the realities of Dutch citizens with different origins in different ways by creating the technicity for technologically mediated racialization. Through my historically comparative perspective, I aim to both denaturalize contemporary categorization practices—by showing that there are many ways of counting people—and contextualize them—by showing that current practices are the result of a long history and many different factors.

For reasons of clarity and structure, I will distinguish between three time periods. First, I will discuss the years 1899–1940, from the foundation of CBS until World War II, when counting by hand was replaced with a process of mechanization. During this time, social stratification in the Netherlands mainly happened on the basis of religion, a societal structure commonly referred to in the Netherlands as “pillarization” (*verzuiling*). The second period, 1945–1980, extending from the end of World War II to the end of a period of labour immigration and decolonization in Suriname and Indonesia, is characterized by a move from mechanized counting to digital counting. This was the period in which the social organization of the Netherlands moved from pillarization to multiculturalism. The final era, from 1980 onwards, has been characterized by a

national discussion on and scholarly investigation into immigrant integration and the multicultural society, combined with the increasing datafication of governmental information. One aspect of this period is relatively stable and particularly relevant for my sociotechnical history: the census has become 'virtual'. This moves the object of the process of counting from people themselves to their data, as it exists in bureaucratic systems. The use of the three different periods described here is mainly intended to structure this chapter; I am not claiming that there are any 'hard breaks' in social or technical developments that determine what happens in these periods. In this historically oriented infrastructural inversion, I will present the results in chronological order to make for smooth reading experience and a better connection with the following chapters.

### **Social Stratification through Pillarization (1899–1940)**

In 1899, CBS was founded as an independent administrative body of the Dutch government (Kuijlaars 1999). Between 1899 and 1940, the institution was able to largely centralize the production of governmental statistics – a practice that had been spread over different government departments and municipalities previously (Kuijlaars 1999, 217–23). In addition to institutional reforms, CBS also initiated technological developments in census counting and statistics. In 1899, CBS still used self-reporting by hand with pen or pencil on a pre-printed form with different paper colours for men, women, and children (den Dulk and van Maarseveen 1999, 352). These forms were then counted by hand by CBS employees. In 1916, CBS was the first Dutch organization to install the Hollerith punch-card system in order to effectively count imports and exports, a previously rather time-consuming activity (Kellenbach 1999a, 121). For the 1920 census, the Hollerith machine was considered, but for unknown reasons, the "classi-compteur", another mechanical counting device, was chosen for this purpose (Kellenbach 1999a, 76).<sup>19</sup> The handwritten census forms were counted by women using this classi-compteur. However, from 1930 onwards, the Hollerith system was chosen for the census. CBS employees handled the census forms, converted to individual punch cards for each person. These cards were then counted automatically with machines at the CBS building (Kellenbach 1999a, 75). The results of the censuses were printed in reports with tables in both Dutch and French.<sup>20</sup> The census that was planned for 1940, was cancelled due to the German occupation of the Netherlands in World War II (Kuijlaars 1999, 227).

---

<sup>19</sup> For a detailed description of the classi-compteur, see Lars Heide (2009, 153–55).

<sup>20</sup> Official governmental documents regularly featured French translations until the 1950s. This was a hangover from Napoleonic times (Van der Burg 2007).

## “Nationality” and “Place of Birth”

Between 1899 and 1930, the population of the Netherlands increased from a little over 5 million people to almost 8 million. The percentage of people considered ‘foreigners’ or ‘aliens’ grew from 1.04 % in 1899 to 2.22% in 1930.<sup>21</sup> This shows that while the numbers of immigrants were relatively low, their portion of the population was slowly growing, as represented by the census reports of 1899, 1909, 1920, and 1930. In this period, the origins of people were counted in two different ways: in terms of *nationaliteit* (nationality; see Figure 5) and of *geboorteplaats* (place of birth; see Figure 6). All nationalities and places of birth were counted individually and their counts were reported in the attachments of the census reports. The tables with the numbers on nationality features the names of the countries that are considered most important in the main body of the report. In 1899, nationalities were not ordered alphabetically in the main report, but by number of people in each category, ranging from almost 32,000 people from Germany to 233 people from Italy (CBS 1901, 3). From 1909 onwards, the list was extended with eleven extra countries. With the exception of Austria–Hungary, a dual monarchy that fell apart in 1918, the top five countries remained the same. As of this point in time, the order of the countries and their clustering seemed to be based roughly on their proximity to the Netherlands and the number of immigrants present. The lay-out of the documents was another decisive factor in the order and number of countries, in that the available space on a printed page and the page break in the middle split the table in two sets of eight columns (see Figure 7 and Figure 8). The first eight countries are neighbouring countries<sup>22</sup> and other European countries from which larger numbers of people migrated to the Netherlands.<sup>23</sup> The following eight columns are reserved for the remaining countries,<sup>24</sup> ordered roughly by number of immigrants.<sup>25</sup> The residual categories remain the same during this period: they contain people from “other countries”, denoting the totals of the non-featured countries that are mentioned in the attachments, stateless people,<sup>26</sup> and people with an “unknown” nationality.

---

<sup>21</sup> In 1899, the Netherlands had 5,104,137 inhabitants of which 52,989 were foreigners (CBS 1901). In 1930, the Netherlands had 7,935,565 inhabitants of which 175,850 were foreigners (CBS 1933).

<sup>22</sup> Germany, Belgium, Great Britain and Ireland, and France.

<sup>23</sup> Italy, Switzerland, Austria, and Hungary.

<sup>24</sup> Russia, the United States, Sweden, Norway, Denmark, Luxembourg, Turkey, and, depending on the year, Transvaal Colony or Poland, and Greece or Czechoslovakia

<sup>25</sup> The only quantitative exception is Poland. It is unclear why it is ranked number 15 in 1920 and 1930 with numbers of people (roughly 2700 in 1920 and 5900 in 1930) that exceed the quantities from all other countries, except Germany, Belgium, and Austria.

<sup>26</sup> Statelessness between 1900 and 1940 was often caused by World War I and other conflicts in pre-World War II Europe.

<b>Census Year</b>	<b>1899</b>	<b>1909</b>	<b>1920</b>	<b>1930</b>
<b>Nationality</b>	The Netherlands	The Netherlands	The Netherlands	The Netherlands
	Germany	Germany	Germany	Germany
	Belgium	Belgium	Belgium	Belgium
	Great-Britain and Ireland	Great-Britain and Ireland	Great-Britain and Ireland	Great-Britain and Ireland
	France	France	France	France
	Austria-Hungary <sup>27</sup>	Italy	Italy	Italy
	Italy	Switzerland	Switzerland	Switzerland
		Austria	Austria	Austria
		Hungary	Hungary	Hungary
		Russia	Russia	Russia
		United States	United States	United States
		Sweden	Sweden	Sweden
		Norway	Norway	Norway
		Denmark	Denmark	Denmark
		Luxembourg	Luxembourg	Luxembourg
	Turkey	Turkey	Turkey	
	Transvaal Colony <sup>28</sup>	Poland	Poland	
	Greece	Czechoslovakia <sup>29</sup>	Czechoslovakia <sup>30</sup>	
<b>Residual Categories</b>	Other countries	Other countries	Other countries	Other countries
	Unknown	Stateless	Stateless	Stateless
	Stateless	Unknown	Unknown	Unknown

Figure 5: "Nationality", compiled from the census reports of 1899, 1909, 1920, and 1930 (CBS 1901; 1911; 1922; 1933).

<sup>27</sup> Austria–Hungary was a dual monarchy that roughly comprised the current territories of Austria, Hungary, Czech Republic, Slovakia, and Bosnia Herzegovina between 1867 and 1918.

<sup>28</sup> Transvaal-colony was a colony of Great Britain between 1877 and 1910 and part of what is now South Africa. Most likely, this category only refers to white former inhabitants, mostly of Dutch or English descent, of Transvaal Colony, since black people were exempt from the right of citizenship (De Villiers 1896, 23–24).

<sup>29</sup> Czechoslovakia roughly comprised the current territories of Czech Republic and Slovakia, between 1918 and 1992.

<sup>30</sup> Idem, see previous footnote.

<b>Census Year</b>	<b>1899</b>	<b>1909</b>	<b>1920</b>	<b>1930</b>
<b>Place of Birth</b>	Current municipality of residence	Current municipality of residence	Current municipality of residence	Current municipality of residence
	Different municipality in same province	Different municipality in same province	Different municipality in same province	Different municipality in same province
	Different province	Different province	Different province	Different province
	In one of the Dutch Colonies	In one of the Dutch Colonies	In one of the Dutch Colonies	Dutch East Indies, Suriname
	A foreign country	A foreign country	A foreign country	A foreign country
	Unknown	Unknown	Unknown	Unknown

Figure 6: "Place of Birth", compiled from the census reports of 1899, 1909, 1920, and 1930. Compiled from CBS (1901; 1911; 1922; 1933).



GEMEENTEN. Communes.	Totaal der werkelijke bevolking. Total de la population de fait.		DE WERKELIJKE BEVOLKING BEHOORDE NAAR DE NATIONALITEIT TOT																	
	M. H.	V. F.	Nederlind. aux Pays-Bas.		Duitschland. à l'Allemagne.		B-Igië. à la Belgique.		Groot-Brittanuie ou Ierland. à la Gr. Bretagne et l'Irlande.		Frankrijk. à la France.		Italië. à l'Italie.		Zwitserland. à la Suisse.		Oostenrijk. à l'Autriche.		Hongarij. à la Hongrie.	
			M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

Figure 7: Excerpt of census report of 1909, featuring part of page 242 with the column that mentions the country names. In the original report, this excerpt appeared next to Figure 7 as the top of the same large table, spread over two pages. Source: CBS (1911, 242–43).

The “place of birth” category in the censuses of 1899–1930 is rather locally defined and is related to the current place of residence. The first two categories in each of these years refer to people who were born and continued to live in the same municipality, or a different municipality within the same province. Given that, before the 1930s, the Netherlands had a large number of municipalities with a population of fewer than 500 inhabitants, this category concerns rather small-scale entities. The third category, “different province in the Kingdom”, likewise refers to places within the Dutch territories. Technically speaking, this should also include the colonies, since they were part of the Kingdom of the Netherlands at the time,<sup>31</sup> however, this is not what is meant here, as the fourth category focuses specifically on those territories, effectively separating them off. Until the census of 1920, the Dutch Antilles, the Dutch East Indies, and Suriname are referred to as “the Dutch colonies”. From 1930 onwards, these localities are instead referred to by their names. “Foreign countries”, which refers to all territories outside of the Netherlands and outside of the Dutch colonies, and “unknown” are the final categories.

### Pillarization and the Racialization of Religious Affiliation

Finally, we need to consider religion as a human characteristic measured in the censuses between 1899 and 1930. From roughly 1850 to 1970, the main form of social stratification in The Netherlands was based on religious affiliations or the lack thereof, a process that is generally referred to as *verzuiling*, or “pillarization”, which can be explained as “as a form of segmental differentiation in a functionally differentiated society, which promotes social exclusiveness and an in-group mentality” (Schrover 2010, 332). This kind of social organization is characterized by

<sup>31</sup> The Dutch Antilles are still part of the Kingdom of the Netherlands as of 2021.

Opinion de fait appartenant d'après la nationalité																				Personen zonder nationaliteit.		Personen waarvan de nationaliteit onbekend.		TOTAAL			
Rusland. à la Russie.		Vor. Staten van N.-Amerika. aux Etats Unis.		Zweden. à la Suède.		Noorwegen. à la Norvège.		Deno- markou. au Dens- mark.		Luxemburg. au Luxem- bourg.		Turkije. à la Turquie.		Transvaal- Kolonie. à la Colonie du Transvaal.		Grieken- land. à la Grèce.		een ander land. à un autre pays.		Personen zonder nationaliteit. Personnes sans nationalité.		Personen van waarvan de nationaliteit onbekend. Personnes dont la nation- nalité n'est pas connue.		Neder- landers. Néerlandais.	Vreem- delingen. Étrangers.		
M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.
H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.	H.	F.
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47		

Figure 8: Excerpt of census report of 1909, featuring part of page 243, with the column that mentions the country names. In the original report, this excerpts appeared next to figure 6 as the top of the same large table, spread over two pages. Source: CBS (1911, 242–43).

voluntary institutional segregation that enables people to only deal with institutions, information, and people subscribing to their belief. Consequently, at the height of pillarization, roughly between 1900 and 1950, it was possible for a Roman Catholic to live his/her whole life within the Catholic pillar. This entailed for instance attending Catholic primary and secondary education, possibly going to a Catholic University, reading a Catholic newspaper, listening to a Catholic radio broadcaster and being a member of the Catholic labour union, etc. (see for example Thurlings 1979). This could all happen in a city or village that simultaneously hosted institutions from several other pillars.<sup>32</sup> Thus, at the end of the 19<sup>th</sup> century and in the first half of the 20<sup>th</sup> century the Netherlands effectively hosted several “pillars” that operated in parallel to each other (M. Hoogenboom and Scholten 2008, 108).

The metaphor of pillarization does not only describe the social isolation of individual pillars, but also how the top of the pillars, the social and economic elite, i.e. religious leaders and politicians, discussed and negotiated shared beliefs, policies, and regulations. The government actively provided the opportunity for religious and ideological groups to start their own organizations by providing funding for religiously differentiated “public” institutions such as schools, hospitals, radio broadcasting organizations etc. In order to make this work, accurate information about the representativeness of the various religious groups was very important. Before WW II, the table describing these numbers was, therefore, a significant part of the census (see Figure 9 and Figure 10). Despite the emancipatory ideals of pillarization, envisioned to give minorities equal rights in providing for their constituents, in practice the different pillars effectively created parallel societies that were hierarchically organized both internally and in relation to each other.

<sup>32</sup> The main pillars in the Netherlands were the Catholics, Orthodox Calvinists, Dutch Reformed, the Socialists, and the Humanist Liberals (Thurlings 1979). It needs to be noted that the humanist liberals were against the principle of pillarization, but felt forced to participate.

Tijdscheen waarin de geboorte plaats had. <i>Groupes d'années de naissance.</i>	K E R K E L J K E													
	Nederlandsch Hervormden. <i>Reformés Néerlandais.</i>		Wetlandsch Hervormden. <i>Reformés Wesleyens.</i>		Remonstranten. <i>Remonstrants.</i>		Christelijk Gereformeerden. (Afgescheidenen). <i>Reformés chrétiens.</i>		Dooptgeen in den. <i>Ambogétoets.</i>		Evangelisch Lutherischen. <i>Luthériens évangélistiques.</i>		Hervormd Lutherischen. <i>Luthériens orthodoxes.</i>	
	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Figure 9: Left side of table with different religious denominations in the Netherlands (CBS 1901, 120)

It has been argued that the sophistication and ubiquity of the larger pillars (i.e. catholic and protestant) in particular, functionally isolated people from the rest of society, hindering mutual understanding and acceptance (Thurlings 1979; M. Hoogenboom and Scholten 2008; Schrover 2010). In addition, the pillars were not created equally. The Catholics and Jews in particular had suffered discrimination and stigmatization in the early 19<sup>th</sup> century, a problem which, despite being granted equal rights by law, continued for decades into the 20<sup>th</sup> century (Homan 1966; Thurlings 1979; Weiner 2015). Pillarization, and its bureaucratic infrastructure, gave the relatively large group of Catholics stability and safety. The much smaller group of Jews, however, did not enjoy such benefits. It has been suggested that the available bureaucratic infrastructure made it relatively easy for the Nazi occupation to find and detain Jews during WW II (Bovenkerk 2000; Weiner 2015, 578–79). 85% of Jewish people in the Netherlands were deported and killed. According to historians, the *combination* of Dutch bureaucracy with a rather advanced railroad network, and specific geopolitical circumstances at the end of the war was one of the main causes of the mass extermination of Jewish people living in the Netherlands (see Croes 2006; de Haan 2010).

These findings seem to confirm the thesis of sociologist and philosopher Zygmunt Bauman, that the racism against, and extermination of, the Jews was not some pre-modern social attitude with new instruments, but rather a racist development “unthinkable without the advancement of modern science, modern technology, and modern forms of state power” (Bauman 2007, 61). It should be noted that I do not discuss the relationship between information systems and the Holocaust to make a historical argument about WW II, since my limited analysis of censuses and secondary literature does not warrant any major conclusions about this topic. Rather, I aim to bring across how we can look at the bureaucratic technologies involved as conditions of possibility—the extermination of the Jews was not primarily or exclusively caused by bureaucracy, but bureaucratic technologies were a necessity for the possibility of the Holocaust to be executed by the Nazi’s in the way and scale that it did.

GEZINDTEN.																	
Behoorende tot de Gereformeerde kerken. <i>Appartenance aux Eglises réformées.</i>		Roomsche Katholieken. <i>Catholiques romains.</i>		Oud Roomschen. <i>Jansénistes.</i>		Nederlandsche Israëlieten. <i>Israëlites Néerlandais.</i>		Portugeesche Israëlieten. <i>Israëlites Portugais.</i>		Personen behoorende tot eenige andere kerkelijke gezindte. <i>Appartenance à une autre culte.</i>		Personen behoorende tot geen kerkelijke gezindte. <i>Personnes sans confession.</i>		Personen van wie de kerkel. gezindte onbekend was. <i>Personnes dont le culte est inconnu.</i>		Totaal der werkelijke bevolking. <i>Total de la population de fait.</i>	
M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.	M.	V.
16	17	18	19	20	21	22	23	24	25	26	28	27	29	30	31	32	33

Figure 10: Right side of different religious denominations in the Netherlands (CBS 1901, 121)

In the period before WW II, pillarization as a type of government did not only structure life within the borders of the territories of the kingdom of the Netherlands in Europe, but also extended to the Dutch colonies. The Dutch government provided the financial means for religious groups to go on missions to colonized areas of the kingdom such as Surinam and Indonesia (see for example Steenbrink 2007; Vernooij 2017). Furthermore, for a long time pillarization remained the organizing principle behind societal institutions in Surinam and Indonesia, such as schools and political parties (van Leeuwen 1968; Gunawan and van den Muijzenberg 1967; Marchand 2014; Lijphart 2018). Taken together, these developments show the rather broad scope and ubiquity of pillarization, which had a profound impact on the organization of Dutch society and its bureaucratic infrastructures, both on the main land and in its colonies, before and after WW II.

## Implications

In the period between 1899 and 1930, the people that were counted had a relatively high measure of autonomy. Since they provided information about themselves, theoretically they could also provide inaccurate information. Who counted as Dutch and who did not was relatively straightforward: people with a Dutch passport were considered Dutch, regardless of their place of birth or the nationality of their parents. Furthermore, the local nature of the categorizations also indicated that CBS considered migration to be not only a transnational but also an intranational phenomenon. Taking into account province and municipality of birth, the reports demonstrate that CBS considered “internal” migration to be a significant statistic as well. It should be noted, that internal migration in the Netherlands in this period was not only measured in the populations of cities and towns but also tied to people—the variable that was measured were people moving from one place to another, accompanied by statistics about cities and towns shrinking or growing.

On an instrumental level, the statistics of CBS about people concerning “nationality” and “place of birth” seemed to be considered in a strictly legal sense. The CBS quantification of religious affiliations was based on self-reporting. Where statistics about “nationality” and “place of birth” mainly reflected the results of intranational migration and international migration due to war and economic circumstances, the information about religious denominations was directly coupled to institutional pillarization and related policies.

On an epistemological level, “nationality” and “place of birth” did not seem to be racialized before WWII. One could argue that it is very likely that the people counted were almost exclusively white, but the kind of use of the concept “nationality”, as well as its function in the produced tables, does not suggest that it was linked to ideas about race or ethnicity. In stark contrast, censuses in colonized parts of the Dutch kingdom intricately linked ideas about race, nationality, and religion at the time (see Stoler 1989, 153–55; Prianti 2018, 106). Information about religion in the investigated Dutch censuses, however, did have racializing connotations given the recent history of this period. Furthermore, tables about religious denominations can be linked to specific policies causing not only protection, but also isolation and segregation of religious minorities. Despite the negative aspects of this particular institutional organization, pillarization would function as a model for the “integration” of immigrants after WW II.

On an ontological level, technologically mediated racialization between 1899 and 1930 happened mainly based on self-reported religious affiliations. Pillarization, combined with the rather detailed administrations of local governments, can be seen as a technicity of race of this era. This technicity produced race through the registration of religious affiliation. The centralized registration of this particular human characteristic, made it available for racial signification in a time of religious persecution. This situation created, in part, a condition of possibility for the genocide of people racialized on the basis of religion, that was realized in WW II.

### **The Invention of “*Autochtoon*” and “*Allochtoon*” (1945–1971)**

On an instrumental level, the period between WW II and the 1970 census can be seen as a transition period in which CBS moved from counting with punch cards to a more centralized way of counting by adding up municipal administrations. After WW II, a census was initiated as soon as possible, because the government needed a detailed picture of the state of affairs in order to rebuild the nation (CBS 1954, 7). During this 1947 census, CBS still made extensive use of the punch-card system (Kellenbach 1999a, 79). For the 1960 census, however, the first electronic statistics machine was used: an IBM 101. By now the punch cards contained an additional technique, called “marked sensing”, which consisted of stripes that could be read by a card punch machine. In 1957, the first computer (an Electrologica X1) was bought for 700.000,- Dutch

guilders<sup>33</sup>, which was subsequently used for the 1960 census (82). Before the 1971 census, a new Philips P1400 computer was bought with 128 kilobytes of memory and three changeable drives of 7,25 megabytes each (84). Machine readable census forms proved error prone, so instead CBS still relied on punch cards. These punch cards were read by an IBM 360/20 that wrote their information on magnet strips, which subsequently could be loaded in the P1400 computer.

Since a small but significant portion of the Dutch population refused to partake in the census, CBS also used municipal registrations to fill in gaps in their data. This development was the starting point for CBS to move from self-reporting via punch cards to a networked form of information collection that completely bypassed the people involved. In a table in the 1971 census report that is important to this research (see Figure 14), the first two columns show the numbers of men and women that were counted through a process of self-reporting on punch-cards; the third and fourth columns show the number of people that were counted “administratively”, using the municipal administration as source. Prior to the 1971 census, public outrage ensued over the fact that, increasingly, statistical tasks of CBS were performed by a computer, deemed a “people-sorting machine”(CBS 2018c). In addition, the government was planning the introduction of a *Burger Service Nummer* (civil service number), with which all Dutch citizens would be continuously registered in the *Centrale Persoonsadministratie* (Central Administration of People) (Kuijlaars 1999, 394–95). This idea antagonized many among the Dutch, eliciting memories of WW II, when the German occupier was grateful for the meticulous Dutch administration. As I already mentioned, among other details this administration contained the addresses and religions of all citizens, including Jews. Resistance to automated administrative systems resulted in a situation whereby a relatively small number of people refused to cooperate with the census of 1971 (see Figure 11). CBS was able to replace their missing counts with counts from the municipal registration database quite accurately.

---

<sup>33</sup> 700.000 guilders would be roughly 2.300.000 euro today when corrected for inflation



Figure 11: Protest against the 1970 census. The sign of the protester in the middle reads: "Refuse the census. Once you have entered the computer, you can never leave!" (The Hague, 10 February 1970). Photo: Rob Mieremet, Collection Anefo/Nationaal Archief.

## “Nationality” and “Place of Birth”

In the 1947 and 1960 census reports, the “place of birth” table once more contains the options “current municipality of residence” and “other municipality in the same province” (see Figure 12). Compared to the censuses of 1899–1930, the locality level here seems to be upscaled a little, because the “different province in the Kingdom” category is split into eleven different categories for each of the provinces in the 1947 census, and into “adjacent province” and “another province (not adjacent)” in the 1960 census. This allows for more detail on migration at the province level. The main text of the 1960 census report explains that most of the people counted as migrated from “foreign countries” are from “Indonesia and other parts of the Empire in the West” (CBS 1966, 13). This suggests that people from the Dutch colonies are now included in this category. Finally, the “unknown” category is not mentioned in the 1960 census, possibly because the numbers are very low, and/or because the category was no longer deemed important.

<b>Census Year</b>	<b>1947</b>	<b>1960</b>	<b>1971</b>
	<b>(table in attachments)</b>	<b>(table in main report)</b>	<b>(main report)</b>
<i>Place of Birth</i>	Current municipality of residence	Current municipality of residence	Dutch
	Other municipality in province (split up in 11 provinces)	Other municipality in the same province	
	Total of the Netherlands (total of categories mentioned above)	Total current province of residence (total of categories mentioned above)	
	East- or West Indies	Adjacent province	
		Another Province (not adjacent)	
	Different countries of residence (Split up in Germany, Belgium, Great Britain, France , Italy, Poland, Austria, and ‘other’.	A foreign country	Non-Dutch
	Unknown	-	-

Figure 12: Place of birth in the censuses of 1947, 1960, and 1970. Compiled from: CBS (1954, 78; 1966; 1978), and Schreven (2004).<sup>34</sup>

<sup>34</sup> Not compiled from the actual census but from Schreven (2004, 154). It is unknown whether this level of detail was reported in the main report.



It is in the main text of the 1947 census that the word “autochtoon” appears for the first time, and in the explanation of the definitions of the 1960 census that the word “allochtoon” emerges for the first time to describe people born in the same municipality and people born elsewhere (different municipality, province or country) (CBS 1954, 42; 1966, 7). The significance of this should not be underestimated, since at the time the meaning of the conceptual pair “autochtoon” and “allochtoon” was 1) not yet the subject of stigmatization, and 2) still very locally defined. In the 1971 census, place of birth was collected and analyzed, but in tables in the main census report the emphasis is reduced to either Dutch or non-Dutch.

The table featuring the nationalities of foreign inhabitants of the Netherlands in the census report of 1947 is ordered by continent and has the title: “Foreigners by nationality” (see

Figure 13). It lists all the European countries in alphabetical order. Then, Africa is mentioned, of which only the South-African Union is highlighted; the other category is “other Africa”. This is followed by the North American countries, the United States and Canada (residual categories here are Central America and South America). The “Asian” countries featured are “China”, “Turkey”, and “Palestine” (with the residual category “other Asia”), which shows that at the time, for Dutch demographers there was no distinction between the Middle-East and South and South-East Asia. From “Oceania”, only “Australia” and “New Zealand” are mentioned specifically, and no residual category is stated. The final two, residual, categories are “unknown” and “stateless”.

In the 1971 census, the table entitled “non-Dutch according to counting method, sex and nationality” arranges countries differently (see Figure 14). First, some of the countries of the former European Community, a predecessor of the European Union, are listed: Belgium, France, Great Britain, Italy, and West Germany. The category of “other countries from the European Community” consists of Denmark, Ireland, and Luxembourg, as mentioned in footnote 2 of the table in Figure 14. Then, so-called *wervingslanden*, or “recruitment countries” are mentioned: Greece, Yugoslavia, Portugal, Spain, and Turkey. The Dutch government targeted these countries to attract workers in an attempt to fill large numbers of vacancies in specific sectors of the economy. This campaign lasted roughly from 1964 to 1970 (Cornelis 1990, 23–24). The African countries mentioned here are Algeria and Tunisia, which are lumped together, and Morocco, again followed by a residual category: “other African countries”. “America” features the United States, Canada, and a category of “other American countries”. From Asia, China and Indonesia are highlighted, followed by “other Asian countries”. Finally, Oceania is not split into individual countries.

## Staat 27. Vreemdelingen naar nationaliteit, 31 Mei 1947

	31 Mei 1947		
	M.	V.	Totaal
1	2	3	4
<b>Europa</b>			
Albanië .....	-	3	3
België .....	12 401	12 680	25 281
Bulgarije .....	7	6	13
Denemarken .....	100	113	213
Duitsland .....	9 822	18 053	27 875
Finland .....	10	11	21
Frankrijk .....	544	1 073	1 617
Griekenland .....	75	71	146
Gr. Britt. en N. Ierland .....	1 172	1 975	3 147
Hongarije .....	406	571	977
Ierse Vrijstaat .....	11	5	16
Italië .....	1 797	1 577	3 374
Joegoslavië .....	567	494	1 061
Luxemburg .....	38	66	104
Noorwegen .....	75	92	167
Oostenrijk .....	834	1 483	2 317
Polen .....	3 108	2 095	5 203
Portugal .....	18	8	26
Roemenië .....	50	46	96
Sovjet Unie 1) .....	145	524	669
Spanje .....	38	47	85
Tsjecho-Slowakije .....	458	525	983
IJsland .....	-	4	4
Zweden .....	74	66	140
Zwitserland .....	443	399	842
<b>Afrika</b>			
Unie van Zuid-Afrika .....	51	55	106
Overig Afrika .....	3	6	9
<b>Amerika</b>			
Verenigde Staten van Amerika..	366	338	704
Canada .....	136	130	266
Midden Amerika .....	22	12	34
Zuid Amerika .....	52	44	96
<b>Azië</b>			
China .....	1 019	251	1 270
Turkije .....	22	10	32
Palestina .....	23	12	35
Overig Azië .....	41	27	68
<b>Oceanië</b>			
Australië .....	8	30	38
Nieuw-Zeeland .....	4	-	4
Onbekende nationaliteit .....	1 036	1 747	2 783
Zonder nationaliteit .....	12 715	11 332	24 047
<b>Totaal aantal vreemdelingen ....</b>	<b>47 691</b>	<b>56 181</b>	<b>103 872</b>

1) W.o. Estland, Letland, Litauen.

Figure 13: Nationalities as featured in the 1947 census. Source: CBS (1954, 49).

Staat 4a. Niet-Nederlanders naar telwijze, geslacht en land van nationaliteit, 28 februari 1971

	Door tellers geteld		Administratief geteld	
	M.	V.	M.	V.
	x 1 000			
<b>EUROPA</b>				
<b>E.G.-landen</b>				
België .....	11,6	11,3	0,4	0,3
Frankrijk .....	1,2	1,5	0,3	0,3
Groot-Brittannië .....	5,5	5,2	1,1	0,6
Italië .....	9,7	5,3	2,0	0,6
Bondsrepubliek Duitsland 1)	17,6	17,5	0,8	0,7
Overige E.G.-landen 2) ....	0,5	0,6	0,1	0,0
<b>Wervingslanden</b>				
Griekenland .....	1,7	0,7	0,9	0,4
Joegoslavië .....	2,4	2,1	2,8	0,5
Portugal .....	1,8	1,4	1,6	0,8
Spanje .....	13,7	5,5	5,8	3,0
Turkije .....	18,2	3,8	7,1	1,3
Oostenrijk .....	1,6	1,4	0,0	0,0
Zwitserland .....	1,2	1,0	0,1	0,1
Overige Europese landen ....	2,6	2,6	0,4	0,4
<b>AFRIKA</b>				
Algerije en Tunesië .....	0,3	0,1	0,1	0,0
Marokko .....	11,4	0,6	9,4	0,2
Overige Afrikaanse landen ...	0,5	0,4	0,2	0,1
<b>AMERIKA</b>				
Canada .....	1,3	1,0	0,1	0,1
Verenigde Staten .....	2,7	2,7	0,6	0,4
Overige Amerikaanse landen ..	0,3	0,4	0,1	0,1
<b>AZIË</b>				
China .....	1,1	0,5	0,5	0,1
Indonesië .....	4,1	3,6	0,5	0,4
Overige Aziatische landen ...	1,3	1,2	0,5	0,4
<b>OCEANIË</b> .....	0,9	0,8	0,0	0,0
<b>GEEN 3)</b> .....	14,1	8,9	0,6	0,5
<b>TOTAAL</b>	<b>127,3</b>	<b>80,1</b>	<b>35,9</b>	<b>11,4</b>

1) Hieronder zijn alle personen met de Duitse nationaliteit begrepen.

2) Denemarken, Ierland, Luxemburg.

3) Incl. de personen die als nationaliteit de "Molukae" hebben opgegeven.

Figure 14: Nationalities as featured in the 1971 census. Source: CBS (1978, 1b: niet-nederlandse nationaliteiten:64).

The order and structure of this list is rather different than it used to be, reflecting a society that is in motion. Turkey is no longer listed as an Asian country, but as a recruitment country. The placement of Algeria, Tunisia, and Morocco as African countries is interesting, since elsewhere in the same report they are listed as “recruitment countries” as well (CBS 1978, 1b: niet-nederlandse nationaliteiten:27). By the time of the census, Indonesia had gained independence, and is, therefore, no longer considered a Dutch colony here, nor does it count here as part of the Dutch territory. The Netherlands was in a process of decolonization and labour migration. It, therefore hosted larger numbers of people with more than one passport.<sup>35</sup> The report, therefore, explicitly mentions that:

people with more than one nationality, one of which is Dutch, are filed under the Netherlands. All other persons with more than one nationality are filed under their first-mentioned nationality. The ‘persons without nationality’ category also includes the people that registered as Maluku Islanders. (CBS 1978)<sup>36</sup>

“First mentioned” in this context means whichever country of which people have a passport is filled in first on the census forms. This is a significant detail, since it shows that Dutchness is defined by nationality, rather than place of birth. The “first-mentioned” nationality still gives some agency to the people counted, except when one of the nationalities is the Netherlands. In addition, it explicitly mentions that people with a double passport of which one is Dutch, are counted as Dutch. It is this rather straightforward categorization practice that is going to change after 1971.

### **Moving from Pillarization towards Multiculturalism**

In addition to the already described public outrage about privacy, there was also debate—that lasted until the end of the nineties—in Dutch society about the status of immigrants, later referred to as the “minorities debate” (Rath 1991). An influential publication that fuelled this debate was *Allochtonen in Nederland* (“Allochtonen in the Netherlands”) by the sociologist Hilda Verweij-Jonker (1971). In her redefinition of “allochtonen” or “foreigners” (“*buitenlanders*”), she mentions three reasons for the choice of groups included in her report. The groups under discussion (1) have come to the Netherlands in large numbers and in a relatively short time, (2) have been subject to some form of government policy to facilitate their arrival in Dutch society, and (3) are clearly physically recognizable because of their skin colour and a language incomprehensible for Dutch people (Verweij-Jonker 1971, 7). The third reason makes explicit the

---

<sup>35</sup> Suriname declared independence in 1975, see Cornelis (1990, 18–20).

<sup>36</sup> People from the Maluku Islands were awaiting the dissolution of Indonesian occupation, which has not happened yet.

racial connotations, by referring to skin colour, and ethnic connotations, by referring to language, that have since stuck to the construct of *allochthon* and the particular groups discussed in the book: repatriates (from Indonesia), Ambonese (Malukan's), Surinamese, Antillians, labour migrants (mostly Turkish and Moroccan), Chinese, refugees, and students from the 'Third World' (mostly African and Asian countries). The discourse surrounding some of these groups would later dominate Dutch debates about integration and "the multicultural society".

Between 1945 and 1971, the *ontzuiling* of Dutch society began (often translated as "depillarization") (Dekker and Ester 1996). The institutionalized social stratification of Dutch society on the basis of religion started to disintegrate. However, simultaneously with the gradual disappearance of the traditional Dutch pillars, new groups that started to settle in the Netherlands were encouraged to organize in a manner that looks suspiciously similar to what used to be known as a pillar. Instead of using categorizations based on religion, the Dutch government addressed newcomers in terms of their nationality. It has, therefore, been argued that new policies that attempted to "integrate" and "assimilate" newcomers in Dutch society were largely based on the rather conservative ideology known as "pillarization", albeit under the guise of "multiculturalism" (M. Hoogenboom and Scholten 2008; Weiner 2015).<sup>37</sup> The growing unease of the Dutch population with a central registration of religious affiliations due to the atrocities committed against Jewish people in WW II, combined with the need to govern new kinds of societal configurations based on national affiliations, encouraged the use of statistics based on nationality. The latter, in turn, started to gain significance at the expense of information about religious affiliations.

On an instrumental level, the definition used to describe person with a Dutch nationality remained the same as in the previous period; anyone with a Dutch passport was considered to be Dutch. The tables containing information about nationality did, however, gain in significance as their content now no longer merely reflected geopolitical developments—that is, the numbers of immigrants from parts of the world in turmoil—, but increasingly started to steer national policies concerning immigrant settlement. At that point in time, a word such as "ethnicity" was, however, not yet part of the vocabulary used by CBS to describe populations in the Netherlands.

On an epistemological level, the concepts "*allochtoon*" and "*autochtoon*" are introduced in the censuses of 1947 and 1960 as neutral terms describing people with a non-Dutch or Dutch

---

<sup>37</sup> A common criticism of rightwing political parties in the 1980s and 1990s of this early period of immigration to the Netherlands is that, like the traditional pillars, immigrant organizations encouraged their members to live in relative isolation from the rest of Dutch society, hindering processes of "integration". However, evidence about the context and process of immigrant settlement in the Netherlands does not support this claim and rather seems to indicate the opposite (see Laan Bouma-Doff 2007; Fajth and Bilgili 2020).

migration background. While these terms survived until 2020, the scale at which they describe migration backgrounds has shifted from people that migrated between municipalities in 1947 and 1960 to people migrating between nations after 1970. During the 1960s, the conceptual pair also picked up racial and ethnic connotations through the ways in which policy and public discourse used the terms, until eventually, around 1970, “autochtoon” and “allochtoon” functioned as the discursive construction of Dutchness and non-Dutchness. From this moment onwards, in government bureaucracy, race-ethnic categories no longer referred to a person’s legal status as a Dutch citizen, but rather to his or her supposed cultural association; “autochtoon” signified a person with two parents born in the Netherlands, and “allochtoon” signified a person with at least one parent not born in the Netherlands. In public discourse, however, autochtoon became a signifier for white people that appear and sound Dutch; autochtoon became its Other, signifying non-white, non-Christian people, that visually and audibly did not fit the norm.

On an ontological level, due to societal developments—such as the Holocaust, decolonization, and postwar labour migration—, the human characteristic used for racial signification changed from a person’s religious affiliation to his or her migration background. Furthermore, on a technical level, CBS started to collect information about migration background by combining the information of municipal administrations. As a result, the basis of the discursive construction of “autochtoon” and “allochtoon” in the census moved from self-reporting to a technical form of information collection that completely bypassed the people involved. As a result, from the 1970s onwards, the social construction of the concepts “autochtoon” and “allochtoon” became a *sociotechnical* construction, creating the conditions for the automation of parts of the processes of counting and statistical calculation. This state of affairs created a technicity of race that eventually led to the production of a myriad of different “objective” statistics about how different race-ethnicities relate to all sorts of social problems. Administrative counting proved to be the condition of possibility for the production of the statistical calculations as answers in search of a problem. “Once you have entered the computer, you can never leave” was one of the popular slogans used in the protests against the 1970 census (see Figure 11). This statement proved to be more accurate than the protesters could have imagined at the time. Their outcry against being entered in the “people sorting machine of state bureaucracy” that could be “easily manipulated” by “evil spirits” (Schreven 2004, 186; my translation) was in effect a protest against the “myth of big data” (Couldry 2017)—a protest against the new datafied rationality dominating non-quantified ways of knowing and governing the social, and the discriminating and racializing conditions of possibility it creates for a society.

## The Datafication of Race-ethnicity (1980–2021)

In response to the rising protest among the general public and the ever-greater ease with which the relatively accurate municipal registers could be accessed, CBS decided to no longer hold a traditional census from 1981 onwards (Kuijlaars 1999, 412). New, computerized systems, which could count and perform calculations very fast in comparison to the machines of 1971 and earlier, and the greater ease of storing information on magnetic tape, made the compilation of statistics more time and cost efficient. However, the policy goals these statistics were related to slowly moved to the background. This process of datafication, understood as the transformation of all aspects of life into quantifiable data (Mayer-Schönberger and Cukier 2013), therefore, not only made information more available, but also disconnected that information from its context. As a result, race-ethnicity as data point, was no longer only available in situations in which it was considered relevant and ethnical, but in all situations about which such data existed. It is this datafication of race-ethnicity, as a condition of possibility, that altered the way in which human characteristics were selected for racial signification, as well as the way in which this data would be appropriated by a variety of societal institutions and actors.

The new technologies of storage and distribution had a big impact on the way in which meaning could be produced about race-ethnic characteristics of people, and the frequency at which this information was made available. CBS could now produce year-by-year statistics, which removed the necessity for a census every ten years. From 1984 onwards, data could be entered electronically with magnet strips, and from 1994 the telephone line and the internet were used to transport data from municipal registrations to CBS, and from CBS to third parties, such as ministries and universities (Kellenbach 1999a, 100–101). In 1995, a selection of statistics was also distributed via CD-ROM and via the internet (Kellenbach 1999a, 101). Starting in 2014, all statistics could be checked through the open data portal present on the website of CBS (Creusen 2014). After this time, CBS increasingly focused on not only presenting their data on the CBS website, but also making its data available for reuse through its Application Programming Interface (API).<sup>38</sup> This interface allows third parties to use CBS data for their own purposes (CBS 2018b). As a result, currently, CBS data no longer function merely as a way of measuring the development of Dutch society and the effects of different governmental policies, but rather as a repository of different statistics that is available to whomever is interested. While I will explore the implications of the datafication of statistics further in Chapters 4 and 5, for now it suffices to state that this development does not only have practical effects, but also implies a different

---

<sup>38</sup> An Application Programming Interface (API) is a standardized protocol that allows different computers and data systems to communicate with each other. In this case, it allows third parties to make use of CBS data also when these data are not on their own servers.

relationship between knowledge, society, and information. Instead of using statistics as a way of measuring relevant developments in society and the effect of specific policies, statistics are now a repository in which, through the use of computation, relevance can be found.

To further understand how these categorization practices developed, it is important to know that, from 1985 onwards, people who had lived in the Netherlands for at least five years could file for naturalization. In addition, children born in the Netherlands gained the right to be registered as Dutch (Heijs 1995, 180–90). In the so-called “minority debate”, these measures are seen as a way of empowering immigrants and granting them the necessary means for full integration. For statisticians, however, this caused a problem, since naturalized immigrants ceased to be visible as a group when Dutchness is defined in terms of one’s legal citizenship status. In the rather heated debate on the position of migrants in the 1980s and 1990s, there was a perceived political need for accurate statistics about migrants concerning employment, welfare status, and crime. Since nationality was no longer a distinctive characteristic, CBS reverted to the birthplace of people’s parents. Due to the use of the municipal registration as the main source of information, rather than self-reporting, it became impossible for people to resist categorizations as they were imposed by CBS based on their data in municipal registers, a situation that continues until today.

CBS determines the origins of people by first looking at whether or not both their parents have the Dutch nationality (see Figure 15). If that is the case, any offspring are considered to have Dutch origins, even if the children were not born in the Netherlands. Offspring of two parents with Dutch origins are considered *autochtoon*. The origins of people considered non-Dutch are, in principle, derived from their country of birth. If this is the Netherlands, the country of birth of the mother is chosen as a person’s origin; however, if the mother was born in the Netherlands, then the country of birth of the father is used as informing the origin of their children (Bovens et al. 2016, 23). Only people with two parents who are born in the Netherlands are considered to be “autochtonen”. This means that in the Netherlands, autochthonous parents can have allochthonous children, but allochthonous parents can only have autochthonous children if they have themselves both been born in the Netherlands. This means that it takes one more generation before the migration background of migrant children is considered to be Dutch.



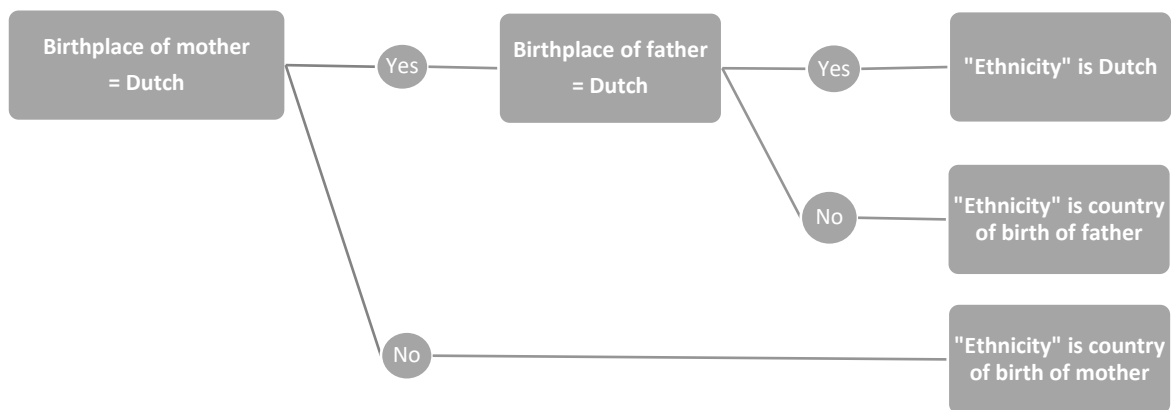


Figure 15: Schematic overview of decision tree used for determining a Dutch citizen's 'ethnic' origins.

Although the grandchildren of people that migrated to the Netherlands are now officially “autochtonen”, CBS has recently started to monitor the “third generation” as well, since, according to CBS, a migration background is “a relevant factor in [one’s] socio-economic development” (CBS 2016). While counting third generation “migrants” – children from parents born in the Netherlands—might seem rather logical and harmless when it is considered a “relevant” factor in socio-economic situations, some issues should be raised though. The classification of third generation Dutch citizens as non-Dutch is not a neutral practice. By producing statistics about third-generation allochthons, CBS is shaping the public discourse in such a way that it seems natural that even the children of children of migrants are not considered Dutch. This resonates with the infamous principle of hypodescent that was used in the United States until late in the 20<sup>th</sup> century, which meant that if a person had only one black ancestor they were considered black, no matter how far back in the family tree the ancestor could be found (Omi 1997, 8). This principle was colloquially referred to as the *one drop rule*, referring to the fact that only one drop of black blood was enough to be classified as black (Denton 1997, 87). The point here is not merely that such a policy is inaccurate, which it is, but rather that it is unethical to keep separating increasingly larger groups of people from discursive Dutchness in an attempt to keep the pool of white “native” Dutch people countable.

From 1989 onwards, a subdivision is made by CBS between “Western”<sup>39</sup> “allochtonen” and “non-Western”<sup>40</sup> “allochtonen” in order to provide more detail about the ways in which different groups of people integrate into Dutch society (Bovens et al. 2021, 9–10). This subdivision was subsequently formalized for policy and government communication in 1999. People from Europe

<sup>39</sup> Westers

<sup>40</sup> Niet-Westers

(excluding Turkey), North America, Oceania, Japan, and Indonesia were considered Western “allochtonen” (see Figure 16). People from Turkey, Africa, Latin America, Asia (excluding Japan), Suriname, the Dutch Antilles, and Aruba are considered non-Western “allochtonen”. While there was no distinction between “foreigners” in the previous periods, now foreign nations are clustered according to their Westernness. People from Suriname, Aruba, and the Dutch Antilles, all considered “non-Western”, have been viewed as problem groups by policymakers, whereas people from Indonesia have supposedly assimilated into Dutch society without any major problems.<sup>41</sup> This perceived difference has severely shaped the image, constructed within Dutch society, of different groups. That the distinction between Western and non-Western countries is not geographically determined suggests that other factors such as race and ethnicity play a role. Hence, making this system the standard for the creation of all major social and economic statistics about Dutch citizens has resulted in the construction of a race-ethnic hierarchy.

### Clustering of Dutch citizens based on country of origin until 2021

Autochtoon	Allochtoon	
<b>The Netherlands</b>	Western	Non-Western
	Europe (but not Turkey)	Turkey
		Africa
	North America	Latin America
	Oceania	
	Japan	Asia (but not Japan)
	Indonesia	[Suriname, Dutch Antilles, Aruba]

Figure 16: Clustering of Dutch Citizens based on country of origin. Source: Yanow and Van der Haar (2013, 240)

### Dutch “Foreigners”

In 2001, for the first time in thirty years, CBS held a register-based census to accommodate requests from the European Union. In the 2001 census report, an English-language document, it becomes particularly clear how problematic it is when Dutch citizens with a migration background are consistently referred to as “foreigners” (Schulte Nordholt, Hartgers, and Gircour 2004). The documents states:

In the Netherlands about eighteen percent of the population can be classified as ‘foreigners’. Foreigners are defined as persons of whom at least one parent is born abroad.

---

<sup>41</sup> It has been argued that the reason for this can be found in the different social and economic circumstances under which these groups settled in the Netherlands. People from Indonesia immigrated in the 1950s, at the time of a big economic boom and with plans to stay indefinitely, whereas people from Suriname immigrated at the beginning of the 1980s, in the middle of a recession (Bovenkerk 1978b, 13–14).

A distinction is made between first and second generation foreigners. A first generation foreigner is a foreign born person with at least one parent born abroad, a second generation foreigner is a person born in the Netherlands with at least one parent born abroad. Foreigners can be further divided into western and non-western; western foreigners come from Europe (excluding Turkey), North America, Indonesia and Japan, and non-western foreigners come from Turkey, Africa, Latin America and the rest of Asia. (Gouweleeuw and Harmsen 2004, 97)

As can be seen in Figure 17, four groups in the “foreigners” category are highlighted: Turks, Moroccans, Surinamese, and Antilleans. These are the populations that have been considered “problematic” since the 1970s – and are, therefore, considered “non-Western”. As noted by Kees Groenendijk:

The typically Dutch distinction between Western and Non-Western allochthones is evidently based on political criteria, namely welfare level, geographical or cultural proximity of the country of origin and assumptions about the problematic character of the group. (Groenendijk 2007, 105)

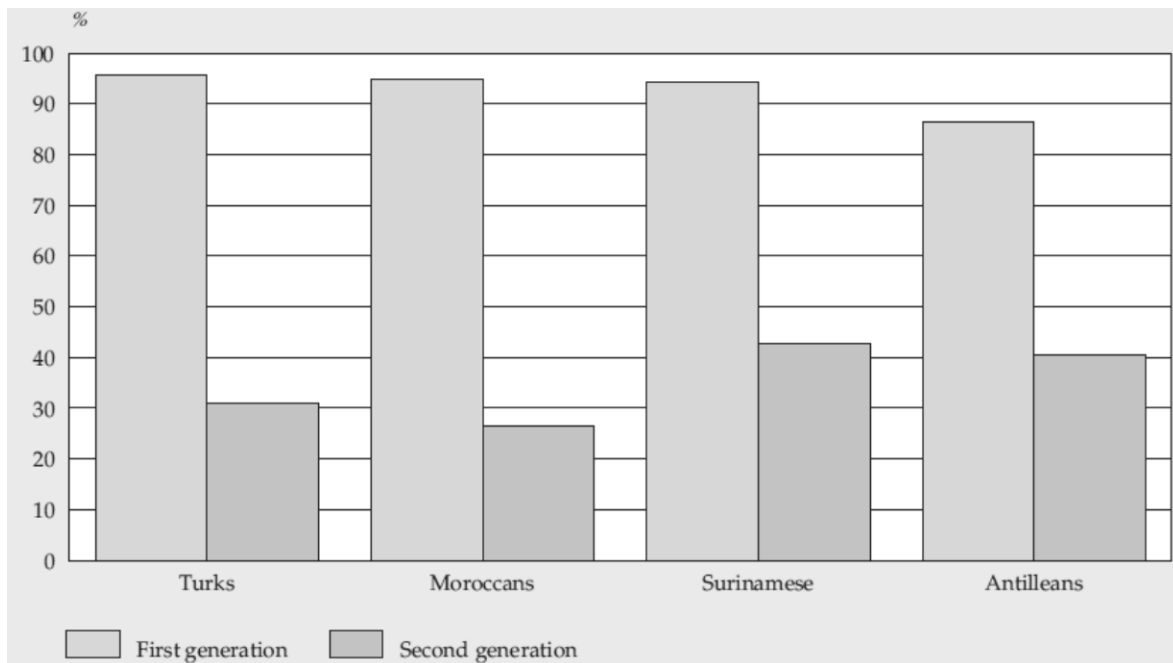


Figure 17: Proportion of the “foreign population” in the potential workforce in the 2001 census. Source: Gouweleeuw and Harmsen (2004, 97).

“The problematic character of the group”, in this context, is an implicit reference to the process of racialization that has occurred with respect to several of the larger immigrant groups in the Netherlands, most notably the Turkish, Moroccan, Surinamese, and Dutch Antilleans. Historians Michael Omi and Howard Winant explain the process of racialization as “the extension of racial meaning to a previously racially unclassified relationship, social practice or group” (Omi and

Winant 2015). In this context, it explains how a personal characteristic, such as having a non-Dutch migration background, has become a socially, politically, and culturally laden idea. In the 2011 census, the duration of stay is included in the table about immigrants (see Figure 18). Here it is interesting that, in the accompanying text, CBS refers to the fact that:

a large majority of immigrants from non-EU/EFTA countries have quite a long migration history. Almost three-quarters stayed in the Netherlands for more than 10 years, and nearly half for longer than 20 years. (CBS 2014, 53)

The message here is that most of the people in this category have been “around” for quite a while. By naming the table “duration of stay”, it is implied, still, that these groups are considered guests and that their stay might not be permanent.

Country of birth	Duration of stay in the Netherlands				
	Total	less than 5 yrs	5-9 yrs	10-19 yrs	20 yrs or longer
	x 1,000	%			
Germany	122.3	20.9	9.8	17.0	52.3
Poland	66.6	60.3	19.0	11.2	9.6
Belgium	50.0	17.5	13.0	18.7	50.8
United Kingdom	47.2	25.3	14.7	21.6	38.4
France	23.4	33.1	14.1	19.5	33.3
Italy	20.8	30.2	12.0	15.8	42.0
Spain	20.4	30.1	12.0	14.0	44.0
Bulgaria	15.0	82.0	8.1	6.8	3.2
Portugal	14.7	34.8	15.6	16.3	33.3
Romania	12.3	53.9	15.4	18.6	12.2
Total	468.1	33.6	13.2	16.8	36.3

Figure 18: “Duration of stay in the Netherlands” in the 2011 census. Source: CBS (2014, 53).

### Dutch Technologically Mediated Racialization

Taking nationality as a self-evident epistemological basis of government policy and the social sciences has been dubbed a form of “methodological nationalism” (Wimmer and Glick Schiller 2002; 2003; Chernilo 2006). Researchers operating within this paradigm “assume that countries are the natural units for comparative studies, equate society with the nation-state, and conflate national interests with the purposes of social science” (Wimmer and Glick Schiller 2003, 576). One of the ways in which the flaws of this epistemology become clear is in the classification of people from areas in the world where borders are contested. The Dutch categorization system

classifies people from Palestine as Israeli<sup>42</sup> and Kurdish Turks as Turkish.<sup>43</sup> People are, therefore, possibly classified as belonging to the very nation for whose prosecution they might have fled. Another pitfall of methodological nationalism is the aforementioned process by which, over time, categorizations become racialized through the assignment of specific connotations to nationality labels. The implementation of methodological nationalism in datafied technologies speeds up this process, as it makes the combination of race-ethnic data with data about crime, poverty, or any other negative markers much easier. As such, this process of technologically mediated racialization cannot be attributed to either technological, social, juridical, or political changes, but rather to the combination of and relation between these factors.

The particular process of technologically mediated racialization as it happens in the Netherlands produces particular normative assumptions. The “autochtoon”/“allochtoon” dichotomy as an epistemology has shaped knowledge production by setting autochthony as the norm. By introducing these terms in its classifications in the 1950s, making them socially meaningful and relevant in the 1960s and 1970s, and making the “allochtoon” and “autochtoon” sociotechnical standards in 1981 with the first virtual census, CBS and Dutch policymakers have created their own, historically and geographically situated, norm against which all ‘Others’ are measured:

Like whiteness, ‘autochthony’ has implicitly (and sometimes explicitly) functioned as the unreflexive norm, a neutral category, a natural fact without a history or relational context. Thus it functions, like whiteness, as a ‘reference category’ (Emirbayer and Desmond 2012; Hartigan 1997) against which deviant cultures can be measured, or a cultural ‘whole’ into which minoritized and racialized Others can be reasonably expected to ‘integrate’.  
(Mepschen 2016, 29)

By using autochthony as a “reference category”, CBS implicitly makes a value judgement about each category in the system, which therefore becomes a hierarchy. Subsequently, through everyday use and in the popular discourse, the dichotomy of allochtoon and autochtoon has become “a racial discourse carried on implicitly in a setting in which the use of the term ‘race’ may be verboten, but where ‘everyone’ knows, and understands, tacitly, the unspoken text” (Yanow and Van der Haar 2013, 250). In the case of racialized discourses imposed on “ethnic” categories in the United States, Dvora Yanow has suggested that the resulting categories can be called *race-ethnic* (Yanow 2003). This critique of the racialization of the labels used has also been

---

<sup>42</sup> In export statistics, however, CBS (2017b) does recognize “areas under Palestinian authority”. Possible reasons for the use of a different classification systems for the movement of goods and money are unclear.

<sup>43</sup> See CBS (2017c).

formulated from a postcolonial perspective (see Wekker 2016) and in terms of the inclusion and exclusion from *moral citizenship* (see Schinkel 2010).

As a result of the pejorative connotations attached to the labels “allochtoon” and “autochtoon”, in 2016, the Netherlands Scientific Council for Government Policy (*Wetenschappelijke Raad voor het Regeringsbeleid* or WRR) advised the government to abolish the “allochtoon” label and replace it with the supposedly more neutral “people with a migration background” (Bovens et al. 2016). The Dutch government subsequently made this change, and as of 2021, most new official policy documents and statistics no longer contain the tainted words “allochtoon” or “autochtoon”. In May 2021, WRR also advised the government to abolish the distinction between the clusters of Western and non-Western migration backgrounds (Bovens et al. 2021). While these are definitely positive developments, it should be noted that these changes are not implemented with the same speed in all parts of the government. For example, until the end of 2021, the categorization system functioned in the same manner for the applications discussed in Chapter 4 – that is, relying on the distinction between Western and non-Western, and with the use of the words “allochtoon” and “autochtoon.” Furthermore, while the government has come up with a new vocabulary and will come up with new ways of clustering people, some parts of the categorization system, such as forced categories with the parents’ places of birth as a basis, are likely to remain the same.

### **“Migration Background” and Colonial Aphasia**

With the increasing institutionalization of the ontology of “migration background”, we have arrived at the state of affairs in 2021. It is significant that, in all the statistics that can be produced, “migration background” merely replaces the “allochtoon”/ “autochtoon” system in terms of functionality. Tables are still laid out in the same way from top to bottom: that is, the total of “native” Dutch, people with a migration background, followed by several specific countries (see Figure 19). Furthermore, as of 2021, “migration background” is still defined in terms of the place of birth of a person’s parents. CBS even explicitly mentions that, while the term “allochtoon” has been replaced by “person with a migration background” at the end of 2016, “the definition was not changed” (CBS 2021b). This creates the rather counterproductive situation of replacing terms that were stigmatized through practices in popular and policy discourse—including statistics—without actually changing the stigmatizing practices.

**Zorgverzekeringswet; wanbetalers met een zorgtoeslag per 31 december**

Gewijzigd op: 9 oktober 2020



Variabelen kunnen geslept worden naar de kop, rijen of kolommen van de tabel. In de kop is maar één item van een variabele te selecteren.

Persoonskenmerken	Wanbetalers met zorgtoeslag							Wanbetalers met zorgtoeslag (relatief)							T.o.v. totaal wanbetalers						
	aantal							T.o.v. bevolking Nederland							%						
	2013	2014	2015	2016	2017	2018	2019**	2013	2014	2015	2016	2017	2018	2019**	2013	2014	2015	2016	2017	2018	2019**
Migratieachtergrond: totaal	191 160	201 410	184 230	162 590	145 170	130 180	117 520	3,7	4,1	4,1	3,6	3,3	2,9	2,7	67,7	67,5	65,3	65,1	63,2	63,2	61,8
Migratieachtergrond: Nederland	96 480	100 670	90 940	80 180	70 580	62 100	54 990	2,4	2,7	2,8	2,4	2,2	1,9	1,8	64,7	64,2	61,6	61,6	59,2	59,1	57,9
Migratieachtergrond: met migratieacht...	94 690	100 740	93 290	82 410	74 590	68 070	62 530	7,6	8,1	7,8	6,7	6,0	5,3	4,9	71,0	71,2	69,4	68,8	67,4	67,4	65,6
Migratieachtergrond: Marokko	13 410	14 390	13 700	12 180	11 270	10 510	9 800	9,4	10,1	9,9	8,6	7,9	7,2	6,7	77,5	78,5	77,0	76,1	75,3	76,2	74,7
Migratieachtergrond: Turkije	11 150	11 760	10 930	9 650	8 700	7 840	7 120	7,1	7,7	7,6	6,6	6,0	5,4	5,0	68,3	69,3	67,4	66,7	64,9	64,5	62,8
Migratieachtergrond: Suriname	18 640	19 180	17 490	15 350	13 710	12 180	10 740	12,7	13,4	12,8	11,1	10,0	8,8	7,9	72,2	72,5	70,0	69,4	68,4	68,2	66,5
Migratieachtergrond: (vm) Ned Ant, Aruba	12 500	12 980	11 860	10 770	9 580	8 580	7 760	19,9	20,7	19,5	17,5	15,4	13,3	11,8	78,5	78,0	75,6	75,6	73,3	73,4	71,8
Migratieachtergrond: overig niet-westers	17 710	19 020	17 780	15 720	14 370	13 420	12 640	6,8	7,1	6,5	5,3	4,6	4,1	3,7	76,3	77,0	76,0	75,2	74,4	74,6	73,3
Migratieachtergrond: westers	21 280	23 420	21 550	18 740	16 960	15 550	14 470	4,4	4,9	4,9	4,2	3,8	3,4	3,2	61,3	60,9	59,2	58,7	56,8	56,4	54,3

Figure 19: CBS table with people who receive health insurance benefits but do not pay for health insurance sorted by migration background. Source: <https://opendata.cbs.nl/#/CBS/nl/dataset/81065ned/table?searchKeywords=migratieachtergrond>.

The dominant ideological assumption in the production of CBS statistics seems to be that migration background is a relevant factor in many societal issues and that reporting on it is a neutral and objective matter. The open data platform of CBS gives users the opportunity to connect the variable “migration background” with a large number of other variables, such as welfare, crime, geographic distribution, and access to healthcare. In addition, all statistics can now also be accessed through an *application programming interface* (API), meaning that people who want to build their own software or website can directly access the necessary data without having to file a request or copy tables by hand. Through this configuration—in which CBS acts as an information source and provides the technical infrastructure—“migration background” functions as an option, a switch, or a piece of code, that can be turned on, or off, at will. This practice effectively nullifies the name change, since for third party applications the name of a table does not matter, the information in it does. In addition, the availability of various statistics causes a disconnect between the information provided and policy goals. Rather than thinking about which information could best be produced to understand a particular societal phenomenon, now the information—a very detailed data set which can be taken apart endlessly—is produced first. Apparent peculiarities in data are quickly mistaken for causal relationships by citizens, journalists, and policy makers alike. In this way, statistics become floating signifiers ready to be cherry-picked to prove this or that statement about society. The currently available infrastructure explicitly affords the reuse of information, decontextualizing and recontextualizing it in the process.

When the name change of “allochtoon” and “autochtoon” is combined with affordances of the available API, the colonial and racial connections of the categorizations are invisibilized. As a result, we witness a rather sophisticated form of colonial aphasia through which the operationalization of race is normalized. Aphasia in this context is meant to describe “the *irretrievability* of a vocabulary, a limited access to it, a simultaneous presence of a thing and its absence, a presence and the misrecognition of it” (Stoler 2016, 157). The words “allochtoon” and “autochtoon” themselves were already disconnecting racial categorizations from Dutch histories concerning genocide and colonization. The new name change actively silences not only Dutch colonial history, but also more contemporary processes of stigmatization and racialization. However, as I will show, a vocabulary change in the context of datafication only makes particular racialized words a thing of the past, while the racializing *function* of the category that is currently described by “person with a migration background” is kept in place. By making this category and all its accompanying data available through an API, it is made more accessible for anyone interested in using this information, while its racializing potential is hidden behind a technological infrastructure. The normalization of reusing the readily available data that is signified by the non-racist sounding “person with a migration background” can be understood as an expression of white innocence (Wekker 2016). The new label and the availability of the data actively hides its racializing function and encourages new uses of the information—several of which will be discussed in Chapters 4 and 5.

On an instrumental level, the categorization based on “migration background” seems to take away most of the racializing connotations that “allochtoon” used to have. However, in the past, categorizations were more obviously related to government policies concerning, for example, pillarization or immigrant “integration”. In the 21<sup>st</sup> century, the production of data is increasingly disconnected from specific policy initiatives and starts to function as a repository in which issues of governance might be found. Data are seen as objective representations of society for which CBS does not always know a purpose. Through their API, other interested parties are actively encouraged to endlessly combine and recombine sets of data. The selection of categories such as “person with a migration background” as well as the specific definitions that are used for categories, therefore, increasingly function in prescriptive ways. Even if societal actors like to draw different lines across Dutch society if a particular problem asks for it, it is made very difficult due to both the technical default, but also because various national social sciences and government departments have to a large extent become accustomed to this way of thinking.

On an epistemological level, race-ethnic classification changed from being based on the nationality mentioned in people’s passports, to the nation of the birthplace of people’s parents, which in some instances is even extended to the birthplace of grandparents. Strictly speaking, this is a rather far-fetched way of determining someone’s migration background compared to the period before the 1970s. From an empiricist viewpoint, one could even argue that this way



of working is less objective compared to simply registering a person's nationality based on their citizenship status. This raises questions about the real purpose of the current, rather elaborate classification system. Taking into account the vocabulary used in the census of 2011, in which Dutch people with a migration background—with a Dutch passport and working in the Netherlands—are unapologetically called “foreigners” suggests that the Dutch government and CBS are consciously and subconsciously trying to keep the category of “Dutch” as racially and ethnically homogeneous as possible. Furthermore, this current classification practice is embedded in a database rather than merely in a registration system; the latter is always contextualized, the former can be entered and perused randomly. This means that data points can be expanded without having to provide a new context; the addition of new data is always in itself a new context that might provide new correlations. The combination of datafication and classification practices that are increasingly based on cultural connotations instead of geographical location and/or citizenship status, creates a situation in which a biologically and culturally essentialist version of Dutchness starts to function as a technical norm. This technical norm is used to monitor, compare, and police Others, while people that already fit the norm by means of their Dutchness get a free pass.

On an ontological level, the increased datafication of government statistics based on race-ethnicity creates a situation in which we can expect more and more government agencies, companies, politicians, and journalists to make use of these data. Datafication of statistics in this context points at the neutralization of race-ethnic categorizations through label changes combined with their incorporation into a technological infrastructure that invites the decontextualized use of race-ethnic data. However, the cultural competence to deal with issues concerning race and ethnicity in a productive, ethical, and antiracist manner seems to be largely missing both from Dutch governmental institutions, and from the Netherlands in general (Brion 2004; Ghorashi 2014; Essed and Hoving 2014b; Özdil 2014; Weiner 2014a; Wekker 2016; Çankaya 2015; 2020). This situation is exemplified by several well-known scandals that involved unethical and/or inappropriate use of race-ethnic information by government agencies. For example, in the so-called *Toeslagenaffaire* (Child care benefits scandal, see (Autoriteit Persoonsgegevens 2020), it became publicly known that the Dutch Tax and Customs Administration had illegally used data on second nationalities and migration background of Dutch citizens in an algorithmic profiling system (Autoriteit Persoonsgegevens 2020, 16). The result of this practice was that around 26000 people, many of which have a migration background, were forced to pay back the subsidies they received for their child care; subsidies they were legally entitled to, based on their income that was below a minimum threshold in the first place. Thus, through the ontological approach to technologically mediated racialization, we can understand the datafication of race-ethnicity as a condition of possibility for racializing and racist conduct by the Dutch state. The ubiquitous availability of race-ethnic information is not viewed here as the main cause for

inappropriate state conduct, but rather an affordance that makes it possible. Technologically mediated racialization in the 21<sup>st</sup> century in the Netherlands has become datafied, networked, and virtual<sup>44</sup>.

## Conclusion

In this chapter I have addressed the first two subsections of this dissertation, which, as I have shown are very much related:

1. What are the origins of the Dutch race-ethnic categorization system and how are its categories defined?
2. How are race-ethnic categorization systems in the Netherlands datafied by government institutions?

Through the infrastructural inversion I have undertaken in this chapter, it becomes clear that definitions and classifications concerning people's origins are not only socially determined, but also heavily rely on the data ontologies through which these definitions and classifications are institutionalized. From 1899 until 1940, categorization was done by relatively time-consuming ways of counting, combined with a rather large amount of agency on the part of the people that were counted. Only a handful of mainly neighbouring countries seem to be of significant importance in the reports. People coming from and living in one of the colonies were still considered Dutch, and foreigners had not been clustered yet in 'Western' and 'non-Western' groups. The period from 1945 to 1980 can be characterized by a move from mechanized counting to digital counting. At the same time, there were significant social developments in terms of decolonization and labour migration. At this point, former colonies became foreign countries, considered Western or non-Western depending on the social position of the relevant groups in Dutch society. To the so-called "recruitment countries" belonged those that will later constitute the category of *allochtoon*. In the case of a double passport, the Dutch passport overrules other passports, which shows that Dutchness prevailed (see Figure 20).

A national privacy discussion combined with increasingly accurate bureaucratic systems caused CBS to start counting people "administratively". The final era, from 1980 onwards, is characterized by a national discussion and scientific investigation of immigrant integration and the multicultural society, coupled with an increasing datafication of governmental information. The increasingly automated means of counting in this period have gone hand in hand with the instatement of a state-imposed and automated nationality category. Clustering nationalities on

---

<sup>44</sup> From the 1980s onwards, all censuses are performed administratively; CBS itself now calls the census a "virtual census" (see Schulte Nordholt, Hartgers, and Gircour 2004).

the basis of who is “recognizable as a group” has turned an otherwise nationality-based categorization system into a race-ethnically informed one. In addition, “administrative” counting in a “virtual” census means that people become the sum of their accumulated data points and data flows. Additionally, now, for people with a second nationality next to the Dutch one, their second nationality determines the assignment of a race-ethnic category. On an instrumental level, the Netherlands has moved from the production of statistics for specific policy goals to statistics as increasingly disconnected data sets that can be interconnected and recreated at will, meaning that people capable of using the open data interface of CBS can select the variables of interest themselves. In this construction, for many societal issues, “migration background” is a default characteristic of which the definition is predetermined by CBS. In this way, race-ethnicity is one of the primary lenses through which people are invited to look at societal issues. Simultaneously, statistics are less often directly connected to specific policy goals, and the policy goals that once motivated their presence are no longer visible; the current infrastructure affords, possibly even encourages, a decontextualized use of statistics.

On an epistemological level, in the censuses of the past century, the Dutch government conceptualized migration on an increasingly larger scale, starting at the municipal level, later at the provincial level, and finally at the national level. Where it is first used to signify any relocation a person went through in their lives, after 1970, it only signified relocations from a different country. Furthermore, straightforward ways of categorization based on nationality were replaced with racializing categories. The rather objective way of counting nationalities clustered on the basis of the continent in which they were located in 1947, was gradually replaced with a categorization system based on the birth places of both parents, which were clustered based on the “cultural and economic similarities” with Dutch society (Keij 2000). Meanwhile, these classifications were decoupled from specific policies and embedded in databases, available to recombine with other data points at will.

On an ontological level, the increased connectivity of government databases create the technicity for technologically mediated racialization to be incorporated in automated policy as well as knowledge producing systems. However, both the label change from “allochtoon”/“autochtoon” to “person with a migration background”, and the possibility of using statistics in a decontextualized manner, sever the signifiers from their racialized past, and racializing present. As a result, the categories referring to the Netherlands’ colonial Others—that is, counting not only people from formerly Dutch colonial countries, but also those people that are racialized because of their cultural, geographic, and economic backgrounds - are caught in a process of colonial aphasia. Whether or not on purpose, this encourages further racializing use of statistics, since people working with data and algorithms will feel that the differently labelled categories no longer work in a stigmatizing and racializing way and are, therefore, more neutral. Furthermore, they are available for calculation and combination in a technical infrastructure that

treats race-ethnic data as just another data point, not unlike age, income, and addresses. This situation creates a technicity of race, that is the condition of possibility, for race-ethnic information to be used in any algorithmic system, and for any purpose in which race-ethnicity is deemed a relevant data point. In the next two chapters, I will take a closer look at three systems that appropriate race-ethnic data for a new purpose, and that are made possible by the CBS data infrastructure discussed in this chapter.

	<b>Instrumental</b>	<b>Epistemological</b>	<b>Ontological</b>
<b>1899 - 1940</b>	Supporting and monitoring the effects of policies related to pillarization.  Mechanical counting	Racialization based on religion.  Migration defined on the municipal level.	Creating the conditions of possibility for surveillance and state violence against particular religious groups.
<b>1947 - 1971</b>	Supporting and monitoring the effects of policies related to migration within the framework of pillarization.  Moving from mechanical counting to digital counting.	Racialization based on nationality, skin colour, and language.  Migration defined more on the scale of provinces.	Creating the conditions of possibility for migrants to start religion based support organizations  Allochtoon/autochtoon becoming informational standards that shape living conditions.
<b>1980 - 2021</b>	Development from monitoring “integration” policy towards a repository in which correlations can be found.  Counting and calculating through datafication.	Racialization based on place of birth of parents.  Migration is defined on the national level.	Creating the conditions of possibility for datafied racialization and discrimination.

Figure 20: Findings in different time periods divided by approach



## Chapter 4: The Appropriation of Race-Ethnic Data in Governmental and Non-Governmental Data Applications

In the previous chapter I explained how, from 1995 onwards, CBS made its statistics accessible through their Statline web interface, and how, after July 2014, all statistics were made available through the CBS Application Programming Interface (API) (Kellenbach 1999b; Welle Donker, Braggaar, and van Loenen 2019, 8). This development made numbers and statistics more easily available for reuse in third party applications. As a result, interested researchers, journalists or citizens no longer had to manually copy or re-enter large amounts of data in a new system. I finished my historical investigation in Chapter 3 with the system of classification practices that was used roughly until the end of 2021 and its dissemination possibilities. In this chapter, I will focus on systems built on top of this available infrastructure and address the third sub-question of this dissertation:

3. How are race-ethnically categorized data appropriated by governmental and non-governmental organizations?

I will, therefore, no longer focus on the relation between classification and datafication, but rather on the relation between datafication and the appropriation of information for new purposes. This allows me to discuss the different ways in which a technicity of race can be, and is, actualized in the Dutch context and how various racializing uses of CBS data should not be understood as excesses, but rather as institutionalized and normalized forms of everyday racializing conduct.

I will discuss the affordances of race-ethnic data in the CBS infrastructure through an investigation of two applications that make use of CBS data: The Allochtonenmeter and the Leefbaarometer. In doing so, I will be able to show how the CBS data infrastructure is not merely providing neutral information, but is functioning as an important part of a racializing assemblage; the CBS API features various technicities of race that are actualized in third party data systems. I will analyze these systems and their institutional embeddedness through a critical data studies and postcolonial studies lens, focusing on the ways in which race-ethnic categories are made available for appropriation through processes of datafication. Datafication, understood as the transformation of all aspects of life into quantifiable data that can be circulated, combined, and recombined endlessly (Mayer-Schönberger and Cukier 2013), leads to a situation in which we can no longer speak of a single system responsible for a single task. Instead, we should understand the governmental data infrastructure combined with its connected systems as a *data assemblage*, which can be defined as “the technological, political, social and economic apparatuses and elements that [constitute and frame] the generation, circulation and deployment of data” (Kitchin and Lauriault 2018, 1). The affordances of this data assemblage are

the result of the interaction of the material properties of technologies with the particular human features selected "for purposes of racial signification" (Omi and Winant 2015, 110) by both CBS and the organizations that work with their data. I will, therefore, discuss both the technological possibilities of the Allochtonenmeter and Leefbaarometer, and the social and cultural context in which these systems are doing their work.

I will consider this social and cultural context through a postcolonial studies lens, in order to highlight and critique the significant role that is assigned to questions of race-ethnicity in the construction of Dutchness. In this dissertation, colonialism is understood not only as the imperial and capitalist project of "bringing wealth and riches to Western nations through the exploitation of others" (McLeod 2010, 9), but even more importantly as "certain ways of seeing, specific modes of understanding the world and one's place in it that assist in justifying the subservience of colonized peoples to the (oft-assumed) 'superior' civilised order of the colonisers" (McLeod 2010, 20). A *postcolonial* perspective, then, highlights patterns of dissonance between colonial legacies and our contemporary postcolonial predicament; by reconnecting ties otherwise lost in translation in dominant and hegemonic cultural narratives, it makes the "invisible visible" (Young 2012, 23). A postcolonial studies perspective on datafication allows me therefore to expose and challenge normative and hegemonic assumptions with regard to factors such as nationality, race, and ethnicity, as well as foreground the situations in which these factors are mobilized in datafied governance.

In analyzing applications of CBS data, I will focus in particular on the indicators used to depict population characteristics and the ways in which they are calculated and visualized. The constructed indicators reflect a transformation of citizens into data indicators, or "measurable types" (Cheney-Lippold 2017). Through datafication and categorization, citizens become objects of politics. These politics produce a population that, while often having been part of Dutch society for over fifty years, continues to be seen, analyzed, and discussed in terms related to migration and "integration," since their identity is continuously constructed in relation to the country of birth of their parents and sometimes even grandparents (Boersma and Schinkel 2015, 1049). I will argue that the performative work of the data systems under scrutiny here is an increasingly important part of the continuous construction of the identities of specific groups of Dutch citizens *as* migrants. As my case-studies also show, this consideration of Dutch citizens as migrants equates with a position outside of society and even with a threat to that society, and is testament to the persistent colonial mindset of Dutch governmental institutions (Wekker 2016).

In this chapter, a postcolonial critical data studies approach is enriched with Harding's (1991) standpoint theory and Haraway's (1988) concept of "situated knowledge", to enable me to reflect on the construction and communication of the position that the Allochtonenmeter and Leefbaarometer take in the production of knowledge about Dutch society. Both Harding and

Haraway stipulate that all knowledge is created from a particular vantage point, emphasizing the entanglement of that knowledge with politics, relevant personal characteristics, life experience and/or cultural background of the actor that produces and communicates it. These circumstances of knowledge production, often referred to as “positionality” or “situatedness”, are, therefore, not merely contextual, but an important part of the knowledge produced. When the positionality of produced knowledge is not accounted for by scientists, it is performing, in the words of Donna Haraway, a “god-trick” (Haraway 1991, 189). Like the data infrastructures and classification schemes discussed in the previous chapter, the Allochtonenmeter and Leefbaarometer exist within a specific national, political, and bureaucratic culture and context, and, therefore, produce knowledge from a particular perspective. Both these applications exist on publicly available websites, and are built on top of the CBS data infrastructure; however, where the Allochtonenmeter only uses CBS data and is connected to alt-right discourse, the Leefbaarometer is hosted on an official government website of the Ministry of the Interior and combines CBS data with information available from other data sources. The different aims of the Allochtonenmeter and the Leefbaarometer allow me to investigate the normative assumptions embedded in datafied knowledge production, and the ways in which these assumptions are communicated to users.

The case of the Allochtonenmeter allows me to question the neutrality of the CBS infrastructure and highlight the political climate in which information about the migration background of Dutch citizens is framed. The particular case of the Leefbaarometer allows me to show the connection between the Dutch governmental data infrastructure and its racialized categories, racializing functionalities of algorithmic applications such as the Leefbaarometer, and government policies such as the Rotterdam Act, a law created to bar people that are not considered “hardworking Dutch people”<sup>45</sup> from social housing. Through this connection, I will be able to show how datafied infrastructures can be considered not only in terms of their more explicit racialized implications, but also in terms of their technicity, as a “condition of possibility” for the development of a racialized, and racializing, data assemblage. Moreover, by investigating two systems built on top of the CBS infrastructure, I will be able show how the label change from “autochtoon” and “allochtoon” to “person with a migration background” affects the meaning making process of technologically mediated racialization in datafied applications.

---

<sup>45</sup> The phrase “hardwerkende Nederlander” or, in English, hardworking Dutch, is a notion often invoked by Dutch rightwing parties in their discursive construction of the Dutch people or “Volk”. This notion simultaneously invokes ideas about who are considered “the people” and who are considered the “Other”, and is used to accentuate how Dutch people with a job live under an undeserved threat of being replaced by migrants (see Mepschen 2012).



I will argue that datafication of race-ethnicity aids in a process of dissociation, in which data becomes discursively disconnected from its colonial and racializing past through datafied practice. Historian Ann Laura Stoler explains the dissociative condition *colonial aphasia* as “the *irretrievability* of a vocabulary, a limited access to it, a simultaneous presence of a thing and its absence, a presence and the misrecognition of it” (Stoler 2016, 157, italics in original). Through the active dismemberment of signifiers such as the words “autochtoon” and “allochtoon” from their racist and colonial past, the power hierarchies they represent can be kept in place. In the case of the much-discussed Dutch cultural icon Black Pete, colonial aphasia manifests as the discursive erasure of history in order to keep a racist symbol with accompanying humiliating practices in place (see Helsloot 2012). In the case of the abolishment of the terms “allochtoon” and “autochtoon”, colonial aphasia is pushed one step further by sacrificing the signifier in order to save the signified; the symbols “allochtoon”, “autochtoon”, “Western”, and “non-Western” are removed and replaced, while the systems, categorizations, and policies that perform racialization remain unaltered, aiding the continuation of the process of technologically mediated racialization.

## Case-study 1: The Allochtonenmeter

On January 25, 2007, the Dutch right-wing website Geenstijl<sup>46</sup> published the Allochtoon-o-meter (see Figure 21), a web application that provided “absolute figures on the number of ethnically Dutch and ethnic minority residents within a -future- postcode” (Prof. Hoxha 2007; translation by author).<sup>47</sup> This web-based application was created in response to public outrage about the practices of a realtor in the Dutch city of Almere. On his website, this realtor had published the percentages of “non-western allochtonen” in neighbourhoods of Almere where houses were on sale. The Dutch National Society of Realtors (*Nederlandse Vereniging van Makelaars*) subsequently named and shamed him for this discriminatory practice (NRC 2009). Provocatively, Geenstijl responded by making an easy-to-use interactive application providing the percentages of “non-Western allochtonen” for all postcodes in the Netherlands, arguing that they were providing a public service (Prof. Hoxha 2007). After all, they argued, they were merely making existing governmental open data, already available at the CBS website, easily accessible to the public.

Geenstijl used a basic design for its no longer existing Allochtoon-o-meter (see Figure 1) which returned the percentage of non-Western immigrants for any particular postal code (Geenstijl 2017). Where the realtor had linked data about the ethnic composition of postcode areas to parameters important in the choice of a home, GeenStijl positioned its app outside the context of real estate and more squarely within the anti-immigrant alt-right discourse. A picture of an apartment building with a lot of satellite dishes dominates the app design, playing into stereotypes associated with neighbourhoods with a large immigrant population. In the Netherlands, this type of dwelling is mainly associated with social housing areas with government-subsidized apartments in lower-income neighbourhoods. The satellite dishes implicitly evoke an image of a population that is not watching television channels commonly available in the Netherlands and that identifies itself more with media from non-European countries. This picture suggests that integration policies are not effective and that immigrants live in their own segregated neighborhoods in the Netherlands.

---

<sup>46</sup> The name of the website Geenstijl.nl literally means “without style”. Their tagline “Tendentieus, ongefundeerd & nodeloos kwetsend” can be understood as “tendentious, unfounded & needlessly hurtful”.

<sup>47</sup> The name “Prof. Hoxha” is a pseudonym of Ambroos Wiegiers, one of the directors of Geenstijl until 2010.

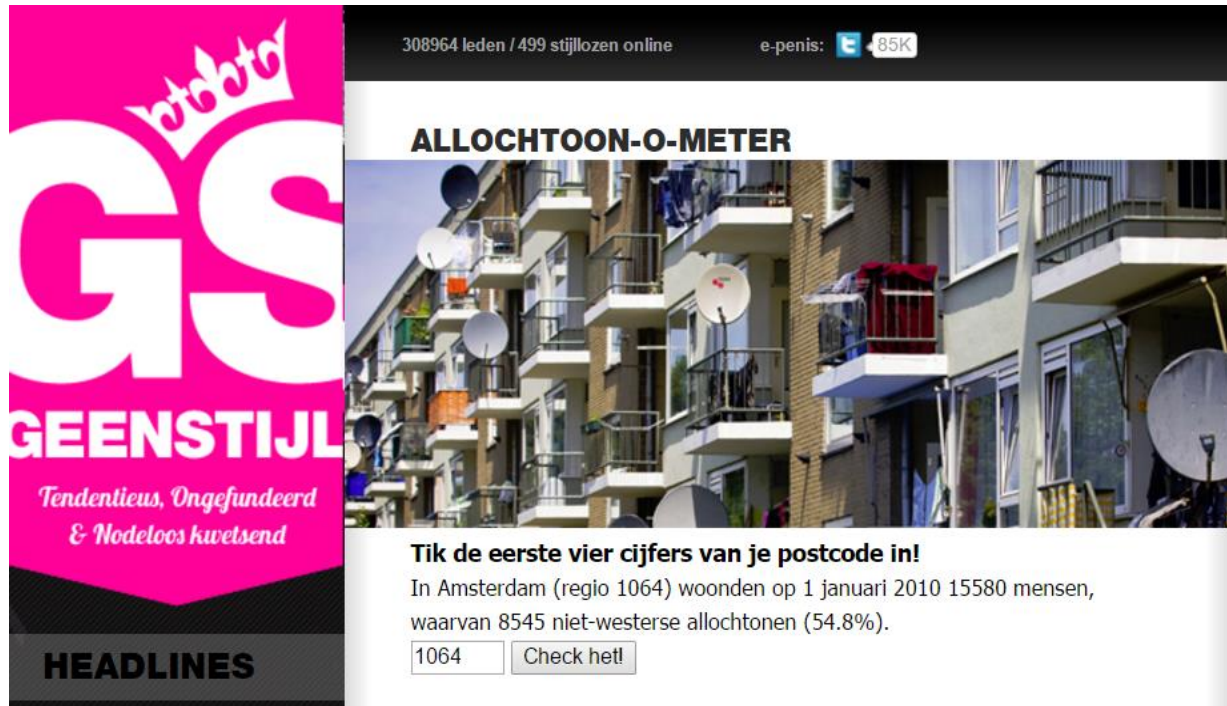


Figure 21: The interface of the Allochtoon-o-meter as shown on the Geenstijl website in 2017. The heading reads “Enter the first four digits of your zip code!” For the purpose of this article, I entered the postal code of a neighbourhood in Amsterdam. The app returns the number of total inhabitants (15580), and the number of “niet-Westerse allochtonen” (8545), as well as the percentage of the total that number is (54,8%).

Although the Allochtoon-o-meter app has been removed now, an almost identical copy was launched with a similar name: the Allochtonenmeter.<sup>48</sup> This app works in a similar way but has a different look, does not seem to be connected to Geenstijl or any other well-known organization, and is hosted on its own URL.<sup>49</sup> Instead of the image of stereotypical housing Geenstijl used, the Allochtonenmeter has had different designs over the years. For example, in 2012 it featured several stereotypical images of nationalistic themes commonly associated with the Netherlands such as pictures of cheese, tulips, and windmills (see Figure 22). This design was replaced around 2017 with a picture of a group of angry-looking cartoon figures wearing what look like niqabs (see Figure 23). This shows how the Dutch cultural Other is not only constructed along race-ethnic lines, but also along religious lines. As of 2021, the picture has been changed to a more diverse group of cartoon figures, still featuring one with a niqab but now flanked by lab coat-clad male characters, signifying the scientific nature of this app (see Figure 24). Furthermore, the only cartoon character signifying a person of colour with an afro haircut literally features a painter's palette with several colors and seems to be either repainting or erasing the seemingly white

<sup>48</sup> See [www.allochtonenmeter.nl](http://www.allochtonenmeter.nl)

<sup>49</sup> [www.allochtonenmeter.nl](http://www.allochtonenmeter.nl) is not explicitly linked to any political party, media outlet, or other organization.

character on the far right of the picture. This can be seen as a graphic depiction of the commonly shared alt-right fear of migrants that “homeopathically dilute”, or simply replace the white European population.<sup>50</sup> The same person of colour character can be seen pressing a button further down the page. This button, which can be used to file a complaint with the website, features the text “if you dare”. This gives the impression that people with a migration background are regularly complaining to this website, but that the owners do not really care.

In terms of usability, the Allochtonenmeter is rather straightforward. After entering a Dutch four-digit postcode, a user can press a prominently placed pink button that reads “get to know your neighbours”. The user is then presented with the total number of inhabitants in the area covered by this postal code, followed by the absolute numbers and percentages of “allochtonen” (total), “Westerse allochtonen”, and “niet-Westerse allochtonen”. The website prominently mentions that these data were gathered by CBS. Such an explicit inclusion of data sources does not only function as a reference but also “performs the rhetorical, persuasive work of ascribing transparency and trustworthiness to visualisations” (Kennedy et al. 2016, 19). Through this practice, the Allochtonenmeter not only literally inherits some of the perceived objective qualities of CBS as an institution, but also communicates objective and neutral values on its interface. Since CBS collects these data every two years by combining the registrations of all municipalities in the Netherlands (see CBS 2021a), the Allochtonenmeter has received an update roughly at the same frequency over the past decade. This practice is reflected in the particular year of the CBS data—in Figure 22, Figure 23, and Figure 24 this is 2010, 2014, and 2020 respectively.

---

<sup>50</sup> Statements like these have been repeatedly made by Thierry Baudet, currently representing the “Forum voor Democratie” in the Dutch parliament. See for example <https://nos.nl/artikel/2326701-baudet-en-vpro-twisten-in-rechtszaal-over-uitzending-buitenhof.html>.

# allochtonenmeter

Het woord allochtoon betekent letterlijk "van een ander land". Tegenover het begrip allochtoon staat de term autochtoon, dat letterlijk "uit hetzelfde land" betekent.

Op dit moment telt Nederland **16.652.323** inwoners waarvan **3.359.603** allochtoon is. Alle cijfers op deze website zijn afkomstig van het CBS (2010).

Toets hieronder de vier cijfers van je postcode in, druk op controleer en je weet wie jouw burens zijn.

POSTCODE:  **CONTROLEER**

## Westerse allochtoon:

Allochtoon met als herkomstsgroepering een van de landen in Europa (exclusief Turkije), Noord-Amerika en Oceanië, of Indonesië of Japan.



Figure 22: Screenshot of the Allochtonenmeter made in 2012

The screenshot shows the 'DE ALLOCHTONENMETER' website. At the top, the title is flanked by two white stars. Below the title is an illustration of several stylized figures in black head coverings, with two larger figures in the center wearing suits. The main heading reads 'Tot uw dienst!' in white on a dark grey background. Below this is a text prompt: 'Toets hieronder de vier cijfers van je postcode in, druk op controleer en je weet wie jouw burens zijn.' This is followed by a white input field. A light blue button with the text 'Meet hoeveel allochtonen!' is positioned below the input field. Underneath the button is a smaller blue button with a thumbs-up icon and the text 'Vind ik leuk'. The bottom section features a dark grey header with the word 'Allochtonen' in white. To the right of this header is a circular icon containing a black silhouette of a rabbit. Below the header, there is explanatory text in Dutch: 'Het woord **allochtoon** betekent letterlijk "van een ander land". Tegenover het begrip allochtoon staat de term autochtoon, wat letterlijk "uit hetzelfde land" betekent.' and 'Op dit moment telt Nederland **16.829.289** inwoners waarvan **3.594.744** allochtonen zijn. Alle cijfers op deze website zijn afkomstig van het **CBS** (gegevens 2014).'

Figure 23: Screenshot of the Allochtonenmeter made in 2018

# ALLOCHTONENMETER

## TOT UW DIENST!



**Toets hieronder de vier cijfers van je postcode in**

Postcode

**LEER JOUW BUREN KENN**



### Huh? Allochtonen? Wat?

Het woord allochtoon betekent letterlijk "van een ander land". Tegenover het begrip allochtoon staat de term autochtoon, wat letterlijk "uit hetzelfde land" betekent (bron: [Wikipedia](#)).

### Hoeveel allochtonen wonen er in Nederland?

Op dit moment (juli 2020) telt Nederland **17.424.761** inwoners waarvan **4.247.258** een migratieachtergrond hebben. In 1996 (eerste jaar cijfers bekend) waren dit er **2.498.715** op **15.493.889**, een stijging van **6,5%**. Alle cijfers op deze website zijn afkomstig van het [CBS](#).

### Westerse allochtoon?

Een Westerse allochtoon is een persoon met als herkomstgroepering een van de landen in Europa (exclusief Turkije), Noord-Amerika en Oceanië, of Indonesië of Japan.

De term 'allochtoon' is eind 2016 vervangen door 'persoon met migratieachtergrond', maar de definitie is niet veranderd.



### Social Media

Op deze website kan je aan de hand van je postcode er achterkomen hoeveel allochtonen er in jouw postcodegebied wonen.

Tweet, like, deel: gun de Allochtonenmeter de bekendheid die zij verdient!

Vind ik leuk

Delen



Figure 24: Screenshot of the Allochtonenmeter made in 2021

In addition to the presented percentages of people, the Allochtonenmeter website contains an explanation of the terms used. It is striking that, at the end of 2021, the tool uses the vocabulary that was abolished by the Dutch government in 2017 for its stigmatizing connotations, as it still talks about “allochtoon”, “Western” and “non-Western.” This is possible because, as a non-governmental website, it is not bound to the officially sanctioned vocabulary. Another factor that allows the Allochtonenmeter to ignore the vocabulary change of the Dutch government is that the categories were only re-labeled, but not differently defined.<sup>51</sup> In a CBS dataset, for example, the percentage of “niet-Westerse allochtonen” in a postcode area before the vocabulary change, would be exactly the same as the percentage of people with a “non-Western” migration background in the same postcode area after the vocabulary change. Furthermore, these numbers could still be found in the same tables, and in the same rows and columns as they would have been before 2017. As a result, no changes had to be made to the Allochtonenmeter website when the Dutch government abolished the terms (see Figure 25). While it could technically be called an algorithm, the code of the website contains not much more than a search function that looks up the appropriate row in a table that combines postcodes with numbers from different clusters of people with a migration background. Its algorithm does not include any calculations since the clusters CBS uses are the same clusters of people used by the Allochtonenmeter.

Pre-2017	Post-2017
“Niet-Westerse Allochtoon”	“Niet-Westerse migratieachtergrond”
“Westerse allochtoon”	“Westerse migratieachtergrond”
“Autochtoon”	“Autochtoon” / “Nederlandse achtergrond”

Figure 25: Equivalents of race-ethnic categorization in old and new vocabulary

Interestingly, the word “autochtonen” and their numbers and percentages are not mentioned anywhere on the Allochtonenmeter. Furthermore, the presence of the niqab makes clear that the Allochtonenmeter is not so much against, for example, German people living in the Netherlands. The white lab coat-wearing men and the absence of numbers on “autochtonen”, communicate the unmarked category in this particular app. It is this unmarked category that produces the perspective in the form of “the gaze that mythically inscribes all the marked bodies, that makes the unmarked category claim the power to see and not be seen, to represent while escaping representation” (Haraway 1988, 581). From this absence, we can trace back the public that was envisioned by the makers and that this application produces: the white autochthonous population of the Netherlands. It is this process that simultaneously produces the cultural Other of this population in the form of the allochtoon. As such, people with a migration background are

---

<sup>51</sup> This is literally mentioned on the Allochtonenmeter website (see Figure 24)



envisioned to be not only non-Dutch but also socioeconomically less advantaged, as well as disconnected from Dutch society at large. Furthermore, by connecting these normative assumptions concerning Dutchness and non-Dutchness to geographical areas, the sentence “find out who your neighbors are” which features prominently on the Allochtonenmeter website, feeds into common fears associated with alt-right migration and livability discourse.

### **Technologically Mediated Racialization in the Allochtonenmeter**

On an instrumental level, it could be argued that the Allochtonenmeter and Allochtoon-o-meter apps are new-right or alt-right political projects that abuse “objective” and “neutrally” collected data. Since they are not owned by a governmental institution, they can use whatever language they please when referring to categories of people with a migration background. Alternatively, they can frame these data however they like, creating new meanings in the process. The fact that CBS did not change the labels, categories, clustering, and definitions, creates an opportunity for third parties to continue to use the abolished idiom without having to worry that their applications produce faulty numbers. The very existence of the Allochtonenmeter and the fact that its functionalities did not have to be revised after the government’s abolishment of the terms “allochtoon” and “autochtoon”, is a testament to how the label change to “people with a migration background” of CBS is merely cosmetic in the context of datafication. In short, the change in terminology from “allochtoon” to “person with a migration background” did not matter in terms of *the racializing meaning and functionality of the data* associated with these terms. In the context of the ongoing datafication of statistics, a simple change of terms, therefore, does not produce any effect.

Consequentially, the epistemological approach to technologically mediated racialization exposes how the allochtonenmeters do not racialize otherwise neutral and objective information; the structure of the information provided by the governmental data infrastructure already—albeit implicitly—contains the racializing meaning that Allochtoon-o-meter and the Allochtonenmeter make more explicit. However, instead of pretending that their information is objective and neutral, both the Allochtoon-o-meter of Geenstijl and the Allochtonenmeter implicitly and explicitly communicate the position from which their knowledge is produced. Where the Allochtoon-o-meter was embedded in the alt-right/populist website Geenstijl, the Allochtonenmeter features several cartoon characters and text that suggest this website is made by and for people on the anti-immigration side of the political spectrum. Problematically, these applications, through their process of framing and representation, contribute to the continuous stigmatization and racialization of migrant groups in the Netherlands. Governmental datasets give them the rhetorical force of objectivity, and provide them with an opportunity to state that they are simply disclosing the facts.

On an ontological level, the functionalities of the Allochtonenmeters are not new, but rather actualizations of a racializing potential that already existed in the governmental data infrastructure. To reiterate, if we understand racial formations always as mediated by specific data technologies, each technology produces a racial ontology based on its historical and sociotechnical context *and* technological possibilities. When we take the Allochtoon-o-meter and Allochtonenmeter not merely as representational instruments, but as actors that perform race-ethnic identities, the purposes of these data systems cannot be seen as unintended by-products of otherwise neutral epistemic methods. Rather, racialization is afforded by an infrastructural system that provides race-ethnically categorized data for reuse. The existence of data that couples race-ethnic information with zip code information in the CBS data ontology can, therefore, be seen as a technicity of race which is merely actualized in the applications discussed in this paragraph. Therefore, the way in which the Allochtonenmeter builds on existing state infrastructure that explicitly affords its racializing project, shows how the provenance of datafied racialization in this context is not incidental or accidental but infrastructural, and, therefore, institutional. To further articulate this argument, I will now discuss how the Dutch government appropriates race-ethnic data in the Leefbaarometer, a system that uses the information about the spatial distribution of people with a migration background in the Netherlands in a much more complex way by combining this information with many more datapoints considered relevant in the calculation of “liveability”.

## Case-study 2: The Leefbaarometer<sup>52</sup>

One year after the launch of the Allochtoon-o-meter, Ella Vogelaar, the Dutch Minister of Housing, Neighbourhoods, and Integration<sup>53</sup> at the Department of Public Housing, Spatial Planning and Environmental Management (VROM) launched the state-run web app called “Leefbaarometer” (Livability Barometer” hereafter LB; see Figure 26). This LB was the result of the integration of two concept systems that were developed earlier: “APK voor wijken” and “Early Warning”. “APK voor wijken”<sup>54</sup> was developed in 2006 by *RIGO Research en Advies* for the Ministry of Public Housing, Spatial Planning and Environmental Management and monitored the social and physical condition of living areas on a small-scale level (Leidelmeijer et al. 2008, 1–2). The development of this instrument was one of the results of discussions with an experimental citizen participation body called “Burgerplatform” of the Ministry of VROM, in which issues of livability were an important topic (see Boogers, Becker, and Tops 2005). “Early Warning”, developed by Atlas voor Gemeenten for the Ministry of the Interior (BZK) was a monitoring instrument of safety and nuisance in neighbourhoods, mainly in cities (Leidelmeijer et al. 2008, 1–2). The decision to combine the aforementioned systems was made top-down, by VROM and BZK, and does not seem to have involved any of the envisioned users of the system that was to be created (Leidelmeijer et al. 2008, 2). The aim of the Leefbaarometer project, then, was to efficiently combine the functionalities of these two systems into one system, feed it with data about all residential areas in the Netherlands, and create possibilities for investigation and monitoring for researchers, civil servants, and interested citizens (Leidelmeijer et al. 2008, 1).

The envisioned aims of the LB were fourfold. First, it aims to facilitate early detection of so-called “livability problems”. In the context of the LB, such livability problems are understood as deficiencies in factors that influence the quality of a person’s living environment (Leidelmeijer and Van Kamp 2003). The second goal is to accurately monitor how the quality of the living environment develops in residential areas that are perceived to have livability problems. Third, the LB aims to provide a first diagnosis about developments in geographic areas by being able to detect “early signals”. These early signals could be, for example, when residential areas still get relatively good scores, but those scores are declining when compared to earlier scores. The final goal is to provide policy evaluations, effect measurements, and “in-depth” research for

---

<sup>52</sup> The information presented in this case study pertains to the Leefbaarometer 2.0, which was replaced with the 3.0 version in March 2022. In this new version, all indicators related to race-ethnicity were removed. This case study, therefore, no longer reflects the current state of affairs, but should be read as a theoretical and historical analysis.

<sup>53</sup> In Dutch: Wonen, Wijken en Integratie

<sup>54</sup> “APK” means *Algemene Periodieke Keuring* (in English: General Periodical Inspection). This abbreviation is mostly known in the Netherlands as the annual mandatory technical inspection of cars, legally required to be allowed to drive on public roads.

policymakers in local governments (Leidelmeijer et al. 2008, 1–2). With these aims and with the aforementioned definition of livability, the LB was created as a publicly accessible interactive system with navigable maps, graphs, and texts that were meant to produce neutral and objective information about the livability of all living areas in the Netherlands.

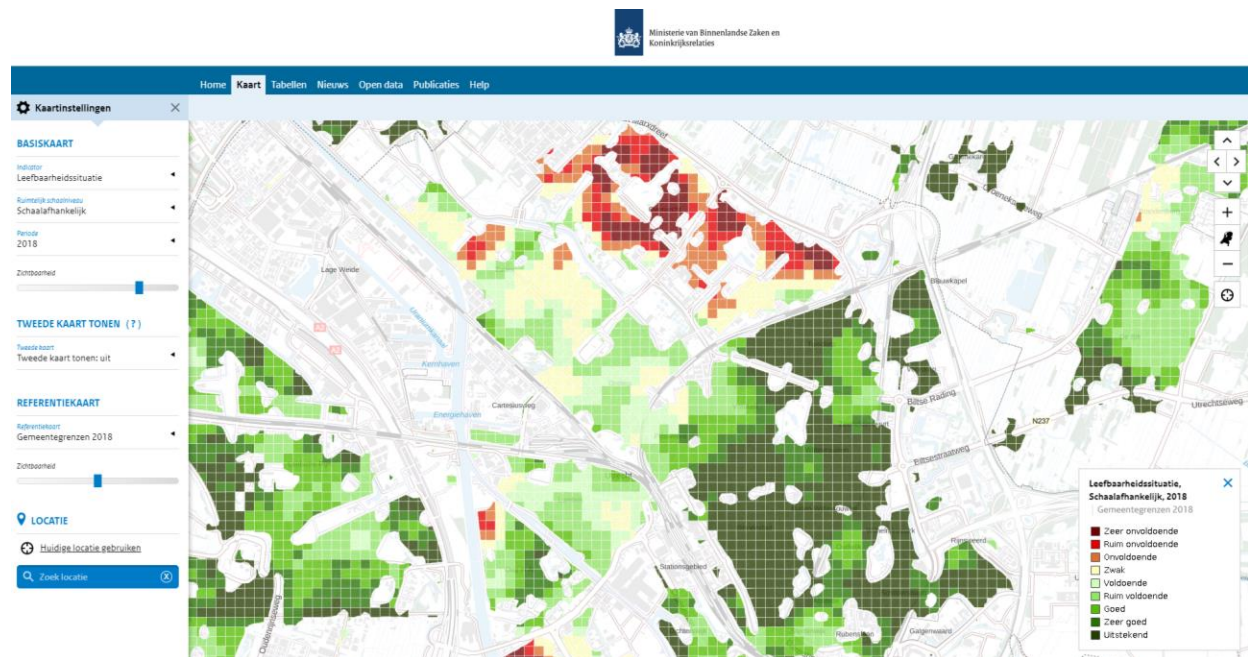


Figure 26: Map of Utrecht, the Netherlands, as shown in the Leefbaarometer app. The colours indicate the level of livability, going from dark green squares, indicating “excellent” living conditions, all the way to dark red, indicating “very insufficient” living conditions.<sup>55</sup>

While the name LB might seem rather straightforward - combining a particular conceptualization of livability (*leefbaarheid* in Dutch) with ‘barometer’ as a neutral indicator of physical/natural humidity conditions - the words are not politically neutral signifiers in the Dutch context. Technically, the meaning of livability in the context of the LB was defined on the basis of a large literature study conducted a few years before the construction of the data system (see Leidelmeijer and Van Kamp 2003). As a result of this study, “livability” was defined as “the measure in which the environment connects to the wishes and demands people have of it” (see Leefbaarometer 2021b, translation by the author). However, the word “leefbaar” also featured prominently in the name of the political party “Leefbaar Nederland” between 2001 and 2006, as

<sup>55</sup> Image acquired in December 2019 from <https://leefbaarometer.nl/kaart/#kaart>.

well as in many Dutch local political parties<sup>56</sup>, the most notable of which are Leefbaar Rotterdam and Leefbaar Den Haag. The most well-known public figure connected to both Leefbaar Nederland and Leefbaar Rotterdam was Pim Fortuyn, a formerly Marxist professor turned into anti-immigration right-wing politician. His anti-immigration stance, as well as his discriminatory antics—he regularly stated that the Netherlands was “full”—were, however, found to be too extreme for Leefbaar Nederland, and its board dismissed Fortuyn from his position as party leader. He remained the leader of Leefbaar Rotterdam and started his own political party Lijst Pim Fortuyn (LPF). Several months later, he was assassinated by an environmental activist, days before the election.

While the political party Leefbaar Nederland was not far right, and the LB officially had nothing to do with this party, both operated within the same discourse on “leefbaarheid”, or livability. This discourse was influenced to a large degree by Fortuyn, and still active successors such as Rita Verdonk and Geert Wilders (Vossen 2010). As noted in the 2008 development report, the developers of the LB viewed the emerging political discourse surrounding livability, and the cooptation of the term livability by political parties, not so much as an unlucky coincidence, but rather as proof of the in significance and relevance of the work they were doing (see Leidelmeijer et al. 2008, 1). As a result, in terms of denotation, livability is simultaneously connected to the physical environment and social services, and historically to the proliferating racial-ethnic discourse of the populist right, particularly, but not exclusively, in the context of local governments.

### Visualizing Livability

It is important to realize that, technically, the LB does not measure but rather predicts a livability score for an area. I will leave a more theoretical and critical discussion of the idea of data prediction for the next chapter on predictive policing. Here I will focus on the race-ethnic data that is appropriated. For its prediction the LB makes use of the results of the WoON-survey, a questionnaire held under 60,000 people living in the Netherlands to investigate the “wishes and needs”(CBS 2021d, translation by author) Dutch people have concerning their living environment. These “subjective” data are combined with “objective” data about the real-estate prices, which, in this case are considered proxies for the popularity of areas and neighbourhoods (Leidelmeijer et al. 2015, 9). The combined “subjective” and “objective” data are compared with statistical data that are available for all residential areas in the Netherlands. The research and

---

<sup>56</sup> Despite the similarities in the names of various local “Leefbaar” parties, there is no nationally organized Leefbaar collective. The local parties bearing this name are, therefore, not officially connected and often have opposing political views.

monitoring affordances of the LB emerge from its interactive geographical map that visualizes a livability score based on a hundred different indicators. By finding correlations between the statistical data and the subjective and objective measurements, a model is created that is then used to extrapolate the measured information to all the residential areas in the Netherlands. This extrapolation, or prediction, is then visualized on a map of the Netherlands. In this way the LB presents a statistical estimate of how livability would be assessed by the average Dutch person in all residential areas in the Netherlands. Since the results are created by extrapolation, the map of the LB should be read not as a geographic map with “facts”, but rather as a weather forecast with a particular degree of uncertainty.

The statistical model of the LB makes calculations on the basis of 100 different of indicators taken from a variety of different data sources. Demographic and housing data are sourced from Bisnode,<sup>57</sup> Centraal Bureau voor de Statistiek (CBS),<sup>58</sup> Rigo,<sup>59</sup> Atlas voor gemeenten,<sup>60</sup> Vastgoedmonitor (VGM),<sup>61</sup> Kadaster,<sup>62</sup> Gemeentelijke Basis Administratie (GBA),<sup>63</sup> and Centrum voor Werk en Inkomen (CWI) (Leidemeijer et al. 2015, 17–21).<sup>64</sup> Indicators taken from these sources are, for example, sociodemographic data, such as the ages and migration backgrounds of people living in an area, real estate data, such as the number of shops within a 1 km radius, the average age and price of houses in a neighbourhood, or environmental data on any parks, forests, or other green areas nearby. Crime and security data are provided by Korps Landelijke

---

<sup>57</sup> Bisnode is an international data science/consulting company originating in Sweden, specialized in helping governmental and corporate actors optimize processes of data-driven decision making.

<sup>58</sup> Centraal Bureau voor de Statistiek, or Statistics Netherlands, is an independent governmental organization that provides socioeconomic data concerning Dutch society; i.e., the data collections consist of data on people’s age, income, and the amount of social benefits in a specified area.

<sup>59</sup> RIGO (Research Instituut Gebouwde Omgeving) is a research and consulting organization with a strong focus on advising semi-governmental actors on how to improve livability levels within municipalities, based on data analytics and research.

<sup>60</sup> Atlas voor gemeenten is a semi-governmental research bureau exploring socioeconomic and cultural phenomena within municipalities in the Netherlands, based on governmental data.

<sup>61</sup> Vastgoedmonitor, or Real-Estate Monitor (author’s translation), is an information system presenting information on real estate issues within the Netherlands. The system connects multiple geo-economic databases to construct graphs and figures concerning real estate phenomena within Dutch municipalities.

<sup>62</sup> Kadaster is a Dutch national real estate registration agency, giving advice on real estate themes such as the value of real estate, property borders, etc.

<sup>63</sup> Gemeentelijke Basis Administratie, or Municipal Administrations Office (author’s translation), is a Dutch governmental administration office registering socioeconomic and demographic data concerning Dutch civilians.

<sup>64</sup> Centrum voor Werk en Inkomen, or Centre for Work and Income (author’s translation), is a Dutch governmental administration office registering all types of socioeconomic data concerning work and income levels within the Netherlands.

Politiediensten (KLPD),<sup>65</sup> and Politiemonitor<sup>66</sup> (Leidemeijer et al. 2015, 17–21). With the help of an algorithmic model, parts of which are under scrutiny in this chapter, these data sources are used to calculate a “livability score” on a scale of a 6ppc (six-number postal code) area with a minimum of 100 inhabitants within that geographically bounded area (see Leefbaarometer 2021a).

Like the choices for particular data mentioned above, the particular design choices for the communication of information should not be considered neutral or logical either. In their 2018 book *Data Feminism*, Catherine D’Ignazio and Lauren Klein explain how we should understand “[a]ny communicating object that reflects choices about the selection and representation of reality” as a “rhetorical object” (D’Ignazio and Klein 2020, 78). Our job as feminist and postcolonial critical data studies scholars, then, is to reflect on the choices made and their effects on the meanings produced in data visualizations. When we take this view to the LB we can see that instead of producing only a table with postal codes and scores, the makers of the LB choose to visualize the measure of livability as an interactive geographical map of the Netherlands (see Figure 26). To make the livability calculation easy to navigate the makers use various rhetorical devices in their presentation.

The first rhetorical device we encounter is the colour coding used by the LB to visualize a livability score. When fewer than the minimum number of forty people live in an area, a part of the map is left blank. This minimal data requirement aims to prevent the accidental identification of individuals in less densely populated areas in the Netherlands. Nine possible colours represent the livability score, ranging from a dark red, meaning very insufficient, to dark green, meaning excellent. On the lowest scale, the default spatial dimension is set to a grid map. The grid map shows blocks of 100 × 100 meters that are coloured according to the calculated livability score (see Figure 28). The size of the squares (100 × 100 m) illustrates that the system is more appropriate for an urban area, since in rural areas the number of people within one square might result in statistical invalidity because of a lower population density. When perceiving the map on a scale larger than the grid map, a complete district is visualized in a colour that resembles the average livability score of that area (see Figure 27 and Figure 28). This can project a skewed presentation of livability within the specified area, as the scores with colours on a lower scale are homogenized. For example, an area including two districts with a livability score of excellent (dark green) that are surrounded by several districts with a livability score of very insufficient (dark red)

---

<sup>65</sup> Korps Landelijke Politiediensten, or Dutch National Police Corps (author’s translation), is the Dutch law enforcement agency on a national level.

<sup>66</sup> Politiemonitor, or Police Monitor (author’s translation) is a Dutch governmental research agency focusing on investigations concerning Dutch criminality and safety.

is represented by a single colour. When the map is viewed on a larger scale, this visualization would imply that the livability score for that entire area is below average, as the colour for that area is presented as red even though those two districts are scored above average. The scale on which the map is visualized is therefore of the utmost importance, as different meanings could be attributed to the visualizations. Clicking on a square shows an information window (see Figure 29), which elaborates on the calculated scores based on the five dimensions with their indicators in comparison with the national average. The national average is a livability score that ranges from amply sufficient to good for all districts within the Netherlands (Leefbaarometer 2019).

In addition, the use of these graphic designs with accompanying colour schemes of shades ranging from green to dark red is signifying a Dutch normative construction of livability. What is, then, silenced in these graphs is that even some of the dark red areas could be considered luxurious, or at least desirable, by people when compared on the basis of alternative terms. For example, for some people, busy city centers, which on average score very high in terms of livability, are the furthest from a desirable place to live. Alternatively, people with a migration background might choose a neighbourhood with many people that share their background over a gentrified neighbourhood that has a higher liveability score. The choice of red and green in this context also signifies safety and/or danger, which is here possibly used in situations where neither is the case. The racializing nature of the Leefbaarometer becomes more apparent in the scores it gives for individual dimensions. As Figure 29 shows, in this application it is possible for areas to receive a negative score for its "*bewoners*", or, in English, "residents". This score is not a proxy for crime or income and does not include factors other than the presence of particular groups in this area, the total of which are apparently valued as "less than average" in relation to the rest of the Netherlands.



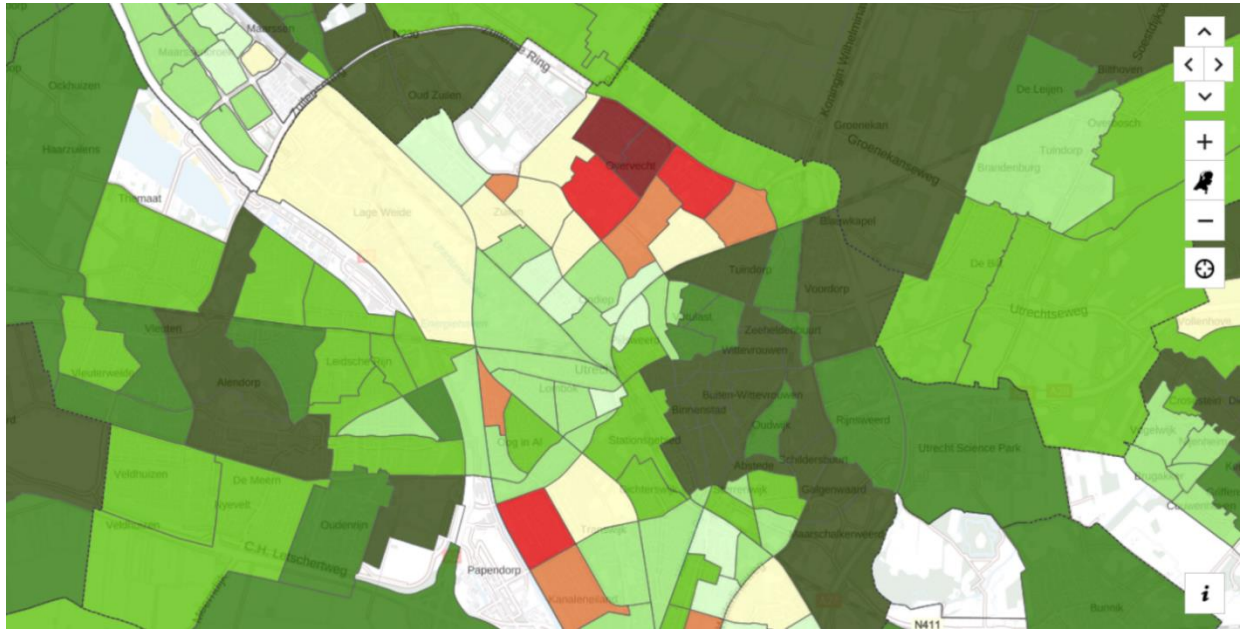


Figure 27: Map of Utrecht, the Netherlands, as shown in the Leefbaarometer app, with the resolution set at the district level<sup>67</sup>

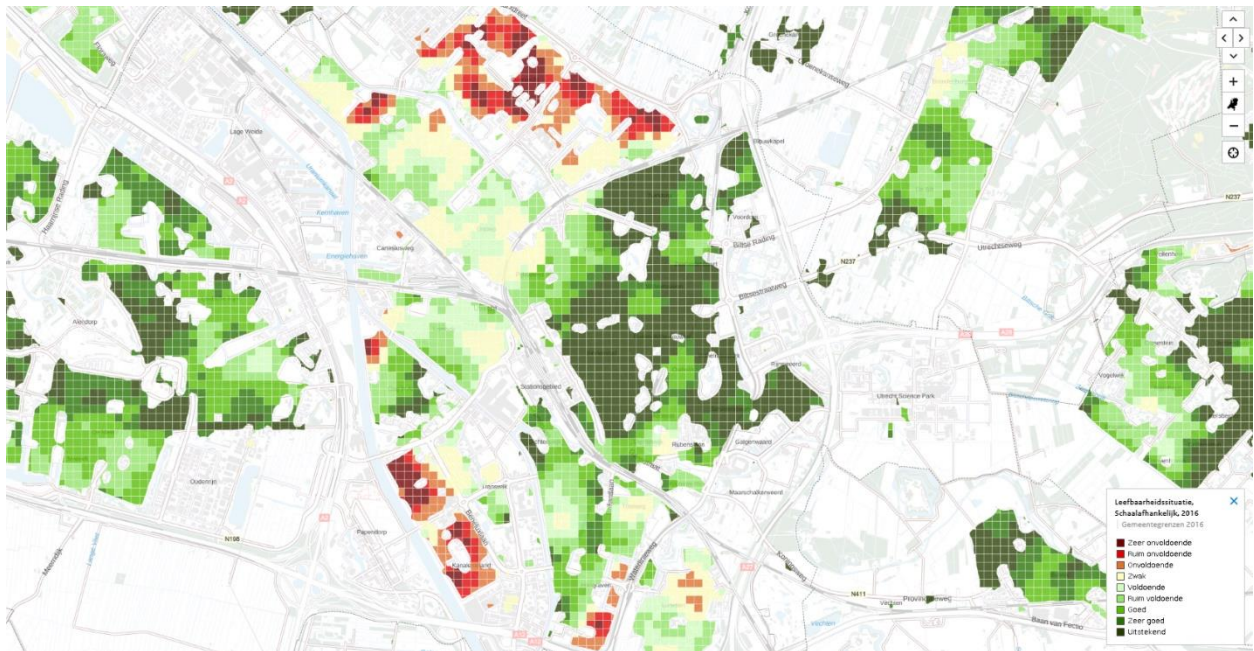


Figure 28: Map of Utrecht, the Netherlands, as shown in the Leefbaarometer app, with the resolution set at the highest level<sup>68</sup>

<sup>67</sup> Image acquired in December 2019 from <https://leefbaarometer.nl/kaart/#kaart>.

<sup>68</sup> Image acquired in December 2019 from <https://leefbaarometer.nl/kaart/#kaart>.

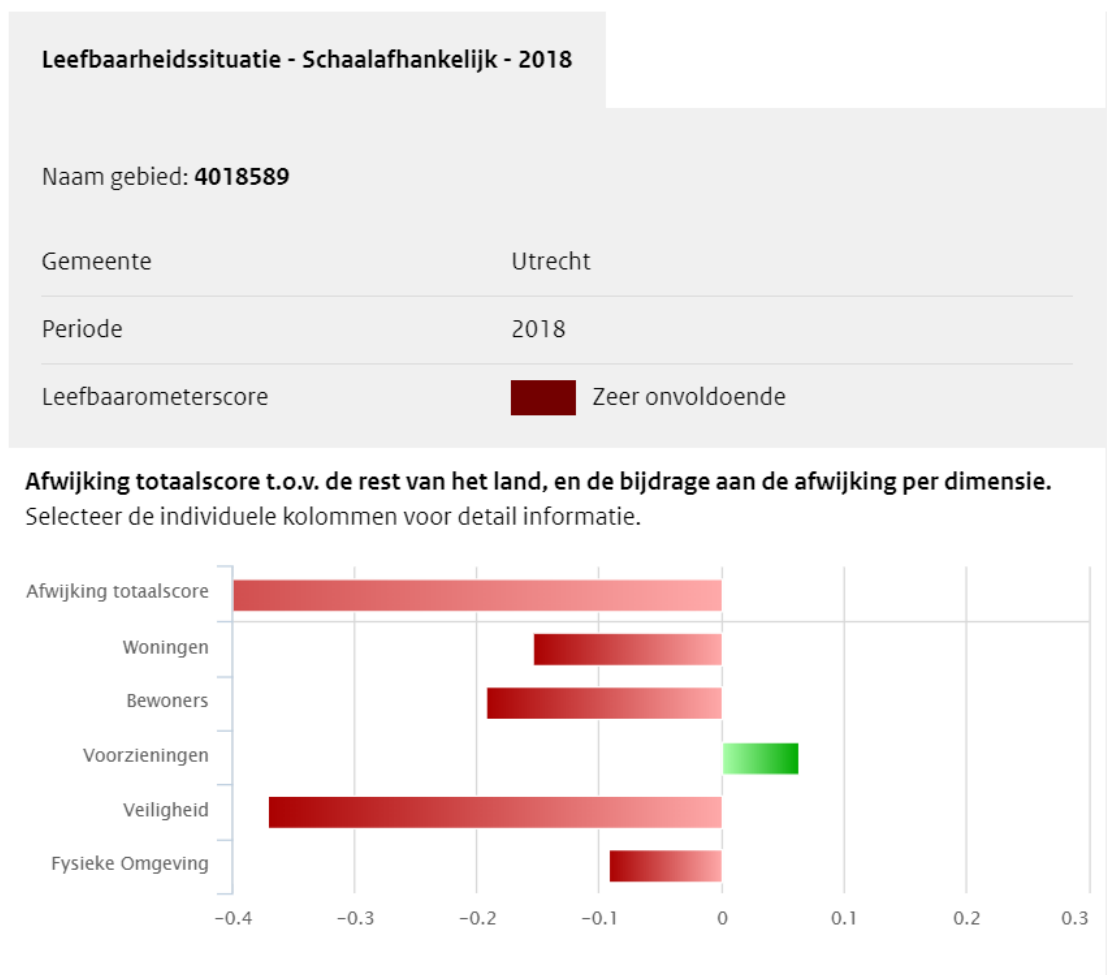


Figure 29: The Leefbaarometer information menu that is visible when a particular area is selected. It shows the municipality ("Gemeente"), the period of the data ("Periode"), the overall livability score ("Leefbaarometerscore"), and a graph showing the deviation of the score from average scores in the Netherlands, both for the total score and the individual dimensions. Note the negative score for the inhabitants ("Bewoners") in this case.<sup>69</sup>

<sup>69</sup> Image acquired in December 2019 from <https://leefbaarometer.nl/kaart/#kaart>.

## Population Indicators

For the calculation of livability, the LB determines the postal code of a certain area and incorporates aggregated data regarding that postal code area in the calculation based on one hundred indicators (Leidelmeijer et al. 2015, 23). As stated in the pie chart in Figure 30, 15 percent of the livability score is determined by the dimension “population.” This dimension consists of 16 indicators (see Figure 31). Indicators 1 to 7 refer to the number of people with various migration backgrounds living in a specified postal code area. Indicators 8 to 10 refer to the composition of the households, incorporating the average number of people in a household, whether or not there are children in the house, and, if there are, one or two parents in the household. Indicators 11 and 12 refer to the number of people who receive social benefits and/or welfare from the state, due to disabilities and unemployment. Indicator 13 is the only indicator that focuses solely on one specific age group, which consists of the elderly (65+). Indicators 14 to 16 refer to the development and fluctuation of people within the age group of 15 to 24 years, composition of households, and people migrating from one postal area to another. For the purpose of this chapter, I will focus on the indicators included in the *bewoners*, or inhabitants, category of the algorithm, as these are indicators that are historically connected with processes of racialization.



Figure 30: Dimensions with percentages in the algorithm used in the calculation of livability<sup>70</sup>

---

<sup>70</sup> Image acquired in December 2019 from <https://www.leefbaarometer.nl/page/help>.

<b>Nr.</b>	<b>Categories</b>	<b>Data Source</b>
1.	Percentage of “Westerse allochtonen”	CBS (GBA)
2.	Percentage of “niet-Westerse allochtonen”	CBS (GBA)
3.	Percentage of Middle and Eastern Europeans	CBS (GBA)
4.	Percentage of Moroccans	CBS (GBA)
5.	Percentage of Surinamese	CBS (GBA)
6.	Percentage of Turks	CBS (GBA)
7.	Other “niet-Westerse allochtonen”	CBS (GBA)
8.	Percentage of one parent households	CBS (GBA)
9.	Percentage of multiple person households (no children)	CBS (GBA)
10.	Percentage of multiple person households (with children)	CBS (GBA)
11.	Percentage of disability benefits recipients	CBS (GBA)
12.	Percentage of social welfare recipients	CBS (GBA)
13.	Percentage of elderly (65+)	CBS (GBA)
14.	Percentage of persons between 15-24 years of age	CBS (GBA)
15.	Development of households	CBS (GBA)
16.	Number of mutations	CBS (GBA)

Figure 31: Indicator table of the “inhabitants dimension” (compiled from *Leefbaarometer 2019*; translation by the author)

Indicators 1 and 2 refer to the categories of Western and non-Western “allochtoon” as a percentage of the whole in a postal code area. There is no category for “autochtoon”. It should be noted that the explanation page of the *Leefbaarometer* website—which prominently features the logo of the Dutch Ministry of the Interior—continued to use the by now abolished stigmatizing word “allochtoon” until the end of 2021. The following indicators are subcategories of this rather broad division. Indicator 3 refers to Middle and Eastern Europeans, which includes the countries of Estonia, Hungary, Latvia, Lithuania, Poland, Slovenia, Slovakia, Czech Republic, Romania, and Bulgaria. To consider these countries as a separate category is in direct opposition to the policy of the European union, which considers the internal movement of people “mobility” rather than “migration” (van Ostaijen and Scholten 2017). Indicators 4, 5, and 6 refer to the three largest groups of people with a non-Dutch migration background: Moroccans, Turkish, and Surinamese. Indicator 7 is a residual category of people with other non-Western origins. For all mentioned populations, larger than average percentages have a negative impact on the population score (Leidelmeijer et al. 2015). It should be mentioned that the factor “migration background” is always working in conjunction with all other indicators in the system, which means that if wealthier labor migrants, colloquially often referred to as “expats”, live in expensive houses in a city center, the score will generally turn out higher. Even in these cases, the factors concerning migration background will lower the scores ever so slightly. This means that, in

general, the very presence of people with a non-Western migration background are considered to have a negative impact on the score of an area.

Another striking aspect of the performance of racialization in the LB is the apparent absence of a category referring to people with a Dutch background. This absence is in line with what other scholars investigating migration and integration discourses have found: “Autochthony” functions as the implicit “reference category” (Emirbayer and Desmond 2012; Boersma 2019). The autochthon/allochthon dichotomy as an epistemology has shaped knowledge production by setting autochthony as the norm against which “deviant cultures can be measured, or a cultural “whole” into which minoritized and racialized others can be reasonably expected to “integrate”” (Mepschen 2016, 29). The absent “reference category” also communicates what the intended publics and/or users are. This was demonstrated earlier with regard to the more right-wing Allochtonenmeter but is also true of the governmentally commissioned Leefbaarometer. This embedded norm of autochthony casts a different light on measurements of livability in terms of “the wishes and demands” *people* have about their environment (Leefbaarometer 2021b). The “people” in this sentence are not so much the total and diverse population of the Netherlands but rather the position of the unmarked category that is taken to be “objective” in the system; the LB is not considered objective merely because it is based on numbers, but rather because the numbers are produced from a middle-class, “autochtoon”, position.

The negative scores associated with people with a migration background and the absence of autochthony as a factor effectively perform state-sanctioned, racialized norms that are apparently rooted in datafied practice. Since the situation in a particular place is always considered in relation to an invisible norm, the percentages of people with a migration background could, in theory, be replaced with the percentages of autochthonous people. In order for the system to produce similar results to the current situation, it would have to systematically overmeasure areas with more people with a Dutch migration background. In practice, it is already giving this bonus by systematically downrating an area with larger percentages of “allochtonen”, while “autochtonen” function as an invisible norm that is not worthy of mentioning anywhere in the system or its documentation at all. By not making explicit these considerations, the datafied application is, quite effectively, obfuscating its racializing assumptions. In the context of the geographic management of the Dutch population, the LB is literally considering people with a migration background to be of a different, lower value, for their living environment than people of Dutch origin.

In the case of indicators about race-ethnicity, we can see quite clearly that people with a Dutch background are not accounted for; this is, however, not the only identity characteristic that is missing in this system. When we look at income sources, we find that indicators 11 and 12 are derived from socioeconomic data concerning recipients of disability benefits, social welfare, and

unemployment data as a percentage of the whole in a postal code area. These indicators generally have a negative value for the livability calculation. The assumption is that income levels lower than the national average automatically result in a decline in livability. There is, again, no indicator for people who are part of the norm in terms of socioeconomic situation. People who do not receive benefits—that is, either people who have a job or people who are not searching for a job—are not represented in the list of indicators. This means that people with jobs, people with spouses with jobs, and people with enough money to not have to search for a job, are considered “normal” in the LB application.

The final norm that can be derived from the indicators of the LB pertaining to the inhabitants of a geographic area, relates to age. Indicator 13 consists of elderly people above the age of 65 as a percentage of a whole in a postal code area. This group is already retired from work or is close to retiring. The indicator is derived from data concerning retirement measures and people receiving a pension. The data is corrected and enriched with data on income levels and generally has a negative value for the livability calculation, as retired people receiving a pension are categorized in a lower income level than people with a job. Indicators 14 and 15 are derived from data concerning the development of people residing in the age group of 15 to 24 years, as these young adults move from their teens to adolescence. There is no indicator for people between the ages of 24 and 65 years, meaning that this range is the norm, a neutral reference, for other ages to be measured against. When we take all the “missing” identities together we can infer that the implied normative identity in the LB is an able-bodied, working, “autochtoon” person between the ages of 25 and 65. In the LB, this identity is considered a neutral identity that does not impact the living conditions of a neighborhood or area. Cultural “Others”, conversely, are made into data objects—dangers that need to be monitored—to protect the wellbeing of data subjects. In its construction of the “autochtoon” working person between 25 and 65 as a neutral category, the LB is configuring its cultural “Other” as an environmental factor, rather than as a data subject with its own “wishes and demands”.

## The Leefbaarometer and the Rotterdam Act as Technological Redlining

By making the LB available through a publicly accessible website, it was envisioned by the Dutch government as a resource for both citizens, journalists, and civil servants enabling them to look into both the current state of affairs and the development of values associated with the livability of residential areas at various levels of scale. The key words here seem to be “potential applications”, or in Dutch “toepassingsmogelijkheden”, as the early development reports do not mention any specific policy, governmental work practice, or existing research difficulties that explicitly suggest a need for a centralized national system like the Leefbaarometer (see Leidelmeijer et al. 2008, 2). The report does frequently mention work that is being done around *probleemwijken* (“problem neighbourhoods”), *grootstedelijke problematiek* (“big city problems”), or *Vogelaarwijken*.<sup>71</sup> However, while the LB seems to fit the task of monitoring the development of “problem neighbourhoods” over time, which specific local government departments, civil servants, or types of researchers, should be using this information, and how, remained somewhat unclear.

In 2014, the algorithmic model used to create the LB map, was recalibrated and reevaluated to adjust to a changing social and technical context. Technical reasons for the changes in the LB are, for example, discontinuation of data collection for particular indicators—which are all provided by other organizations the LB depends on—or when new data points considered relevant become available (Leidelmeijer et al. 2015, 7). Social reasons to include new indicators are 1) the overall changing attitudes of Dutch people towards their respective environments, and 2) geopolitical events such as the inclusion of several Eastern European countries in the EU, creating the possibility for new groups of migrants, previously not captured by available data categories, to settle in the Netherlands (idem). As a result, the indicators of LB 1.0 were not only recalibrated, meaning that their relative importance in the algorithm was adjusted to changing cultural attitudes, but a few new indicators—such as the number of nearby windmills, the chance of earthquakes<sup>72</sup>, and the percentage of Middle and East European labour migrants in an area—were added to the 2.0 model (Leidelmeijer et al. 2015, 11–12). From the 2014 reevaluation it became clear that civil servants of smaller municipalities wanted to know in more detail why

---

<sup>71</sup> In 2007, Ella Vogelaar selected 40 neighbourhoods that were considered to be in decline and that were facing a relatively large amount of social problems. These neighbourhoods were subsequently named after the minister and came to be known as *Vogelaarwijken* (see Sekuur 2010 for a discussion on *Vogelaarwijken* and their media representations).

<sup>72</sup> The number of windmills is growing substantially in the Netherlands due to the need for more sustainable energy sources. The number and severity of earthquakes is growing mainly in the province of Groningen due to decades of gas drilling; this situation is increasingly causing distress and a depreciation of the value of houses in this area (see de Kam and Idsardi 2019).

areas received particular scores. After all, a score does not give any indication about what needs to be improved. Civil servants of bigger municipalities, often larger cities that have their own information systems, explained that they used the LB in conjunction with other information sources (Leidemeijer et al. 2015, 99–106).<sup>73</sup> While not all civil servants are equally positive about its usability—mainly due to LB results not reflecting the experiences and knowledge they have about their jurisdictions—, the evaluation notes clearly show that in many municipalities civil servants not only know but also use the LB.

In practice the LB is not only used to monitor the development of neighbourhoods, but also functions as a measuring instrument to invoke the so-called Rotterdam Act (see Uitermark, Hochstenbach, and van Gent 2017). This law allows municipalities to restrict people with a criminal record, or people on welfare, from access to social housing in areas considered to have a low livability index. This particular measure invoked by the Rotterdam Act can be seen as an attempt to prevent the living conditions in an area from declining further. To clarify, while this measure does not directly discriminate against people with a migration background, the inclusion of race-ethnic indicators in the LB causes neighbourhoods with a high number of people with a migration background to have a higher likelihood of being assigned a low livability score. As a result, they are more likely to be subjected to the Rotterdam Act. Ironically, in the 2015 evaluation of the LB, its creators make clear that they are aware that negative scores for the “population dimension” might create the “perverse policy incentive” (Leidemeijer et al. 2015, 82; translation by the author) of trying to replace the people that are associated with a lower score. However, although the makers called such policy choices “objectionable”<sup>74</sup>, banning particular marginalized people from “problem neighbourhoods” has become one of the main tangible uses of the LB in government policy.

The practice of restricting access to social housing on the basis of algorithmically processed data about migration backgrounds and welfare level, can be understood as a form of *technological redlining* (Noble 2018). The term “redlining”, was used in the USA to designate a process in which credit or other, mostly financial, services are denied to “a subgroup of people in a particular geographic area for reasons unrelated to their creditworthiness” (Squires and Woodruff 2019, 1), which, in practice, was mostly aimed at neighbourhoods with large racial and ethnic minority populations. By invoking the algorithmic LB, Dutch local governments can carry out the prejudiced racializing practice of redlining, while simultaneously covering this practice in an

---

<sup>73</sup> Many larger cities in the Netherlands have their own interfaces or dashboards measuring similar phenomena. See for example Amsterdam (Research, Information and Statistics Amsterdam 2021), Rotterdam (Gemeente Rotterdam 2021), and Utrecht (Gemeente Utrecht 2021).

<sup>74</sup> In Dutch: *aanvechtbaar* (Leidemeijer et al. 2015, 82)



objective and technologically advanced sheen. By specifically selecting underprivileged areas to carry out redlining, municipalities invoking the Rotterdam Act are effectively committing an act of “algorithmic violence” (Onuoha 2018), which, in the context of urban spatial planning, can be understood as a “repetitive and standardized form of violence that contributes to the racialization of space and spatialization of poverty” (Safransky 2020, 1). The combination of the Rotterdam Act, a law that does its work in an already racialized discourse on “probleemwijken”, and the inclusion of indicators related to migration background in the algorithm of the LB, can together be understood as a racializing assemblage that enables technological redlining and regular acts of algorithmic violence.

Apart from the ideological critique on the LB, I would like to make a final point on the basis of the results of the Rotterdam Act. In an evaluation report commissioned by the Dutch Ministry of the Interior, Hochstenbach, Uitermark en Van Gent (2015, xii) conclude that the Rotterdam Act does not to improve the livability of neighbourhoods. It does, however, severely hamper the chances of already marginalized people, who are mostly people with a migration background, to find a suitable home (Hochstenbach, Uitermark, and van Gent 2015, iv). Peculiarly, these findings have not stopped the continuing implementation of the Rotterdam Act in new areas (see Jordan 2018; van Gent, Hochstenbach, and Uitermark 2018). On an instrumental level, we can, therefore, conclude that the net effect of the Rotterdam Act is not an improvement in livability, but rather a diminishing chance for people with a migration background to find social housing. The other side of this coin is that social housing in “problem neighbourhoods” has become more available for people that belong to reference categories of the LB, due to the eliminated competition for housing. As a result, the LB has become rather instrumental in a housing policy that systematically discriminates people with a migration background, while simultaneously privileging “autochtonen” with a job in an increasingly difficult housing market.

### **Technologically Mediated Racialization in the Leefbaarometer**

On an instrumental level, the LB application uses explicit racialized categories, such as “allochtoon” and “autochtoon” in its calculations. As a result of its use of racialized categories, the LB is only able to produce racialized results. Like the Allochtonenmeter, the LB was not affected when CBS changed its vocabulary from the “allochtoon”/“autochtoon” dichotomy to the more neutral sounding “person with a migration background”. In contrast to most other government communications, the Ministry of the Interior did not update the vocabulary on the LB website, at least until March 2022. This shows not only that vocabulary was not so much a priority for the LB, but also that an implemented vocabulary change in an infrastructural system such as that of CBS does not have much effect when the definitions of the categories and the associated data are not changed.

On an epistemological level, and in stark contrast with the Allochtonenmeter, the LB does not explicitly communicate its perspective by providing an overly explicit framing of populations. On the contrary, it almost shines with objectivity, featuring a clean and well-designed map with the logo of the Dutch government featured prominently above its webpage. Implicitly, however, it does communicate a norm of people with a Dutch migration background with a job. Additionally, where the Allochtonenmeter merely produces a number or percentage of people while implicitly assuming that fewer would be better, the LB quite explicitly attaches a value judgement to the prediction it makes: the larger the number of people with a migration background, the more intensive the accompanying colour red. While the evaluation report mentions that civil servants should not read this as an encouragement for “perverse policy incentives” intended to reduce the number of people with a migration background in attempt to get a higher livability index (Leidelmeijer et al. 2015, 82), the reality of the map and its connection to the Rotterdam Act shows that exactly this use is enabled by the combination of the law and the data technology (Uitermark, Hochstenbach, and van Gent 2017).

On an ontological level, the way in which the LB builds on existing CBS data that explicitly affords its racializing project, shows again how processes of datafication enable the obfuscation of embedded racism in the Dutch governmental data assemblage. Furthermore, while the LB was not invented for the Rotterdam Act, its connection with this policy is not accidental. The very inclusion of migration background as a factor in the LB communicates that ethnicity, race, and nationality are considered structuring principles of Dutch society. The negative values associated with these indicators already foreshadow a possible use as a redlining instrument as it incentivizes initiatives that somehow attract people that score higher in the algorithm—i.e. people with a Dutch background and with a job. The Rotterdam Act with its associated policies can, therefore, be seen as a technicity—an explicit potential direction of development—that was already present in the governmental data infrastructure, and even more so in the LB. As such, the Rotterdam Act is an actualization of a racializing potential that already existed in the governmental data infrastructure and the LB. This shows how the categorizations of people with a migration background and the way that they are embedded in the Dutch governmental data infrastructure are not merely epistemic tools, but rather potential world creating entities that can, when enacted in data systems, contribute to unfolding processes of racial segregation in Dutch society.

## Conclusion

In this chapter, I have discussed how race-ethnically categorized data made available by CBS is appropriated by both the Allochtonenmeter and the LB. The performative work of both these systems enable the selection and essentialization of one human characteristic of Dutch citizens in particular: their migration background. The consideration of Dutch citizens as migrants, which equates to being turned into a societal “Other” as my two case-studies show, is a testament to

the persistent colonial mindset of the Dutch government (Wekker 2016). I have shown that the design of the Allochtonenmeter is built on racializing processes that are deeply embedded in institutional and infrastructural information practices governed by the Dutch state. As such, the affordance of the API of CBS to appropriate race-ethnic data is not a politically neutral option but rather the very material infrastructure through which race-ethnicity is made. With the case of the LB, then, I have shown how the Dutch state itself has similar politics, which are effectively obfuscated through the practice of “objectivity” through data and visualizations, normalizing racist ideologies in the process. Through this case, it becomes clear that the technical infrastructure providing race-ethnically categorized data combined with the systems build on top are inherently racializing assemblages, regardless of their intended purposes. More precisely, both systems translate a similar ethno-nationalist ideology into a technologically mediated racializing practice.

Furthermore, the datafied nature of the used information in the LB enables the dissociation of the neutral looking map from the racialized data this map is based on. This technological form of *colonial aphasia* is combined with the use of the words “person with a migration background”, as replacements of “allochtoon”. Both the new vocabulary and the practice of technological redlining (Noble 2018) are haunted by the colonizing and racializing history of governmental bureaucracy. By investigating, not the only the meaning of words, but also the history and performativity of the data categories and technical infrastructure to which they refer, I have made visible how the racializing referent “remains present as ghostly matter” (Dixon-Román 2017, 56). Racializing and racist outcomes should, therefore, not be considered surprising but rather to be expected.

With these case studies I have argued for the use of critical frameworks through which the embedded normative assumptions of datafied systems and epistemologies can be made visible. With the proposed postcolonial critical data studies perspective, it is possible not only to critically engage with the technical apparatus that reproduces inequality but also to place these systems, their ontologies, and their epistemologies in relation to a situation of coloniality. I have shown how already racialized social and cultural data about people with a migration background are mediated by a diverse set of technical systems, taking up new racial meanings on their path. With the notion of the “data assemblage,” I have argued that the construction of racial hierarchies does not happen in one particular location. Instead, datafied racialization emerges through an interplay of a diverse set of actors, which do not require the notion of “race” itself to be already present in the system. Even, or maybe especially, in a culture of racial denial, the process of racialization can happen to a virtually unlimited set of human characteristics. With this understanding, “white innocence” (Wekker 2016) is not only an exclusively human condition but can be carried out through technical systems as well. I have shown how the race-ethnic conceptual pair of *autochtoon* and *allochtoon* are therefore social, as well as *sociotechnical*

constructs, as their meanings rely heavily on the data ontologies through which they are institutionalized. With a performative notion of racialization, this technical institutionalization and infrastructuralization of race-ethnicity can hardly be seen as a neutral and objective epistemic method. In the words of Joshua Scannell, we will have to push past “the insufficient critique that such systems *run the risk* of reproducing racial inequalities. Rather, producing racialized oppression *is all that they can do*” (Scannell 2019, 113); when a datafied system such as the LB uses race-ethnic and socioeconomic data of Others to calculate “livability”, stigmatization is not an accidental byproduct but rather an inherent feature of its design.



## Chapter 5: Race-ethnicity and the Datafication of Policing

In the previous chapter, I investigated two publicly accessible systems, the Allochtonenmeter and the Leefbaarometer, which were built on top of the data infrastructure of *Centraal Bureau voor Statistiek* (Statistics Netherlands, from now on CBS). Neither of these systems are integral to work practices of companies or governmental organizations; the Allochtonenmeter was not made for policy or governance purposes, and the Leefbaarometer is one of multiple tools available to civil servants of local municipalities. Furthermore, the results of the Leefbaarometer only have an indirect effect on the availability of social housing for people living in areas that are deemed less “livable” via the Rotterdam Act. After all, for the Rotterdam Act to take effect, it needs to be approved through a democratic process involving civil servants in several parts of the government. A process that, arguably, has several checks and balances in place. The increasing datafication of governmental processes, however, does allow for the further integration of data applications in the day-to-day practices of governmental organizations (see Eubanks 2018; Brayne 2021). One of the areas in which these data and algorithms are increasingly incorporated is that of policing. In this chapter, I will, therefore, focus on a policing application that differs from the previously discussed systems, in that it does not feature publicly available results, and has direct effects on daily work practice of police officers.

The system under discussion in this chapter is the *Criminaliteitsanticipatiesysteem* (Crime Anticipation System, from now on CAS). It was developed by the Dutch National Police Corps (NPC) for the practice of what is often called *predictive policing*. While there is no consensus about the definition of predictive policing, it can be broadly understood as

the collection and analysis of data about previous crimes for identification and statistical prediction of individuals or geospatial areas with an increased probability of criminal activity to help developing policing intervention and prevention strategies and tactics. (A. Meijer and Wessels 2019, 3)

CAS is one of the latest technological developments implemented by the NPC that fits this definition as it aids in the prediction of petty crimes such as robberies, burglaries, and theft. While it is tempting to understand predictive policing merely in terms of its technological components, its implementation required a rethinking of the processes and purposes of the police organization as a whole.

The NPC underwent two, for this dissertation, notable transformations in the past two decades. First, the formerly locally organized police stations, with their own infrastructures, were organizationally merged into one national police force. Second, the police started to implement a variant of so-called “intelligence-led policing” (ILP), which emphasizes resource management based on information and crime prevention, rather than just crime response. These two

transformations together led to the integration of many formerly local systems into one extensive national database and the development of many new data systems to support the police's tasks. Where in countries such as the United States and the United Kingdom commercially available predictive policing software became popular (see Andrejevic, Dencik, and Treré 2020; Brayne 2021), the Dutch police developed CAS inhouse. In CAS, demographic information from CBS is enriched with available confidential context-specific information from police databases. This system is then used in an attempt to distribute manpower efficiently in terms of timing and location. Since CAS does not predict which individuals are more likely to commit crime, but rather when and where crime will be committed, CAS is considered a “spatiotemporal” predictive system. However, even though CAS is not interested in individuals, it does consider characteristics of people in an area in order to calculate risk. As I will argue, this process can be understood to select specific human characteristics for “racial signification” (Omi and Winant 2015).

Early versions of CAS made use of information regarding the race-ethnic makeup of neighbourhoods; however, in later versions race-ethnicity was no longer an explicit factor. By investigating CAS as yet another system that is built on top of the CBS data infrastructure, but without explicit race-ethnic categories, this chapter allows me to go beyond the third sub-question of this dissertation:

How are race-ethnically categorized data appropriated by governmental and non-governmental organizations?

As I will show, a profiling system like CAS, does not need race-ethnic information in order to perform a process of technologically mediated racialization. The case study, therefore, allows me to think through the practices and consequences of the datafication of race-ethnicity beyond the information that is explicitly mentioned in a system or the policies that surround a system. In the current assemblage of which CAS is a part, the “absent presence of race” (M’charek, Schramm, and Skinner 2014) creates the conditions of possibility for technologically mediated racialization, while simultaneously making processes of racial and ethnic profiling difficult to address.

In order to understand how the embeddedness of a data application in work practice and organizational culture affects technologically mediated racialization, I will consider the policing practice that includes CAS as a “data assemblage” (Iliadis and Russo 2016; Kitchin and Lauriault 2018). To reiterate: the term data assemblage can be understood as “a complex socio-technical system, composed of many apparatuses and elements that are thoroughly entwined, whose central concern is the production of data” (Kitchin and Lauriault 2018, as cited in; Iliadis and Russo 2016, 6). In an analysis of a data system, the focus should be on “the technological, political, social and economic apparatuses and elements that constitute and frame the generation, circulation, and deployment of data” (Iliadis and Russo 2016, 1), in order “to track the ways in

which data are generated, curated, and how they permeate and exert power on all manner of forms of life” (Iliadis and Russo 2016, 2). In this chapter, I will, therefore, operationalize this notion by not only focusing on the data, but also by paying attention to the organizational context in which CAS is put to work. I would like to reiterate that the term “data assemblage” in this chapter does not merely point to CAS as a data system, but more importantly to the sociotechnical network of the Dutch Police that incorporates CAS. The data assemblage that includes CAS depends on categorization practices at CBS, and on the interaction of CAS with the organizational culture and practice of the Dutch police.

In the first part of this chapter, I will rely on an infrastructural inversion and discursive analysis of the CAS system as described in police reports, evaluations, and scientific literature, most notably the evaluation report titled *Predictive Policing: Lessen voor de Toekomst* (Mali, Bronkhorst-Giesen, and Den Hengst 2017). This report describes an extensive test of CAS, which did not result in convincing proof of the efficiency and accuracy of the workflow with CAS. In addition, I will use a list of the data used in CAS, provided in 2017 by Dick Willems, the chief data scientist and creative brain behind CAS, to account for changes made to CAS after the previously mentioned report in 2017. The advantage of the inhouse development of CAS is that the system's algorithms, data, and visualizations can be accessed and adjusted, something that is more difficult to do with proprietary, off-the-shelf systems. In the initial implementation phase of CAS, information regarding what data was used and how was readily available. However, after the national roll-out of CAS in the spring of 2017, no more new tests were performed, and the unconvincing results described in the previously mentioned evaluation report are still the most up-to-date information available about the inner workings of CAS.

In the second part of this chapter, I will combine my analysis of the CAS data system with recent scholarly work and evaluation reports of the Dutch police on issues of racism, ethnic profiling, and organizational culture. Most notably, I will discuss the work of the anthropologists Sinan Çankaya and Paul Mutsaers, who have both shown that the day to day decision-making process of police officers involves similar assumptions about race, class, hotspots, and hot times as the ones incorporated in CAS (see Çankaya 2012; Mutsaers 2015). Taking these results into consideration, I will argue that the iteration of CAS that was used in the previously mentioned evaluation report was a perfect example of the datafication of policing practices, which included forms of technologically mediated racialization. Furthermore, I will argue that the removal of explicit references to race-ethnicity in datafied police practice, while simultaneously expanding the discretionary powers of police officers in an organization that struggles to fight racism within its own ranks, makes race into a tacit presence that is hard to account for in the institutional structures provided. The data assemblage constituted by the combination of CAS, police officers, the organizational culture of the Dutch police, and Dutch culture in general, therefore, generates a system that has only limited safeguards against discriminatory practices.



In the third and final part I will reflect on the problem of accountability when it comes to the implementation of the CAS system, as well as the prevention of ethnic profiling in the Dutch police organization. Information about the working practice with CAS or the issue of ethnic profiling has not been structurally measured or empirically studied on a regular basis by the Dutch police. Rather than merely concluding that limited information about ethnic profiling means that we need to do more research, I will consider the lack of data concerning these issues as a telltale about institutional power relations. The lack of information on ethnic profiling suggests that, until very recently, it was considered not sufficiently important to the Dutch national police to monitor this issue in a structural fashion. The lack of benchmarks and available information makes it difficult to gather how CAS, as part of a larger data assemblage, performs in terms of efficiency and in terms of equality. The lack of a performance baseline makes it currently impossible to know whether updates to CAS or the working practice with CAS can be considered improvements. In turn, this hiatus in knowledge concerning the role of race-ethnicity in intelligence led policing severely hampers academic research into as well as democratic control over the Dutch police organization. A crucial issue that we, therefore, need to address in political debates about the tasks of the police is how this organization can improve its accountability when it comes to bringing into practice the principle of equality, as laid out in Article 1 of the Dutch constitution.<sup>75</sup>

### **Case Study 3: The “Crime Anticipation System” of the Dutch national police**

On 12 December 2012, the Dutch Police implemented a new way of organizing their work process called *Intelligencegestuurd Politiewerk* (Intelligence-Led Policing, from now on ILP) (Kop and Klerks 2009). ILP focuses on the implementation of technologies for the collection and operationalization of information about criminal activity for the management of material and personnel of the police force (Perry 2013; Mantello 2016). ILP is a development that started in the UK and the US at the end of the last century. It was significantly sped up after the 9/11 attacks due to increased technological possibilities and a perceived urgency to predict crime ahead of time (Beck 2009). The attractiveness of datafied solutions to the management of policing lies in its perceived potential to be accurate, cheap, and, therefore, very efficient (Lum and Koper 2015, 260). When looking at ILP within police organizations, ILP is often understood as a “force

---

<sup>75</sup> The translated text of Article 1 reads: “All persons in the Netherlands shall be treated equally in equal circumstances. Discrimination on the grounds of religion, belief, political opinion, race or sex or on any other grounds whatsoever shall not be permitted”.

multiplier” because – in a time of budget cuts and technological development – *intelligence* is perceived as a cheap way of optimizing the policing practices of police forces (Stroshine 2015, 230). From a perspective of police practices, together with other policing strategies, ILP has been a mobilizer to make police forces work in a more proactive manner. It is suggested that crime can be reduced by targeting underlying problems upfront instead of waiting until a crime has been committed or a report has been made (Nix 2015, 275–88). In this fashion, the ideal modus operandi of the police force evolves from reactive to proactive policing (Brayne 2021, 58–60). Such proactive policing implies ways to forecast or predict future criminal behaviour and creates a need for a category of technologies and practices colloquially referred to as “predictive policing”.

The implementation of ILP and its “predictive policing” technologies in the Dutch context was met with some resistance. An often-heard critique during testing was that intelligence should have a more supporting role instead of a leading one (Meesters, Kortekaas, and Tragter 1999, 421; Mali, Bronkhorst-Giesen, and Den Hengst 2017, 41; Mali 2020). It was, therefore, no surprise that in a pilot Dutch police officers felt that CAS interfered with their own agency in making decisions in their daily work practice and had significant doubts about how such a system could support them in their rather diverse set of tasks (Drenth and van Steden 2017). These sentiments are seen as exemplary of the Dutch police culture, which is said to place much value on the expertise and experience of police officers themselves. In an effort to encourage more positive connotations and influence the willingness of police personnel to adopt ILP, the term ILP was translated as *intelligencegestuurd politiewerk*, or “intelligence-driven police work” instead of the more active “intelligence-led policing” in order to imply more human agency (Kop and Klerks 2009, 4). In their “doctrine ILP,” police researchers recognize that leaving room for police expertise to enrich data and results obtained by data-driven systems not only results in more comprehensive intelligence but also helps in persuading reluctant police officers of the benefits of the new approach.

In parallel, the formerly regionally organized police teams were merged into one centrally organized National Police Corps Netherlands (NPC) (Strategische Beleidsgroep Intelligence 2008). This was necessary for the effective implementation of the Dutch variation of ILP. This merger enabled a centrally organized national police database, *basisvoorziening informatie* (basic information provision, from now on BVI) by mandating standards for the collection, aggregation, and storage of operational and crime-related data. Taken together, the ILP doctrine and the centralized formation of the Dutch National Police Force created the conditions for a more extensive datafication of policing practice. In the middle of these developments, which happened roughly between 2000 and 2010, the Dutch police started some experiments by extending systems that were only monitoring certain phenomena, such as numbers of burglaries per neighbourhood, with forecasting possibilities (Mali, Bronkhorst-Giesen, and Den Hengst

2017). These developments together eventually led to the formation of the “Criminaliteitsanticipatiesysteem”(Crime Anticipation System or CAS), which is seen as a logical development within the larger context of ILP as implemented by the NPC (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 16).

CAS is envisioned as the heart of the Dutch approach to ILP, and its goal is both to predict more at-risk areas in a city and, as a result, help the police to work more efficiently (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 91). Despite non-conclusive results both in terms of efficiency (see Mali, Bronkhorst-Giesen, and Den Hengst 2017; Mali 2020) and in terms of user-experience of police officers, CAS was made available to all police teams in The Netherlands in May 2017. At the end of 2017, at least 110 base teams out of a total of 167 in 6 out of 10 districts had incorporated CAS in their working routine (Willems 2017). It could, therefore, be seen as a significant factor in determining where and how the Dutch police would perform their surveillance and patrols.

### Data and Infrastructures

The workflow of Intelligence Led Policing, as practiced by the NPC, is envisioned as a cyclical process (see Figure 32) in which “signals” are captured in “data”, “data” is provided with a context and turned into “information”, which in turn, when combined with other information, can be interpreted as “knowledge” (Kop and Klerks 2009, 8–9). When the available “knowledge” is combined with experience and operationalized into a plan, it is considered “intelligence” (ibidem). “Intelligence” leads to “action”, which in turn results in new “signals” (ibidem). Given the cyclical nature of the process, the analysis could be started anywhere in the cycle. However, because of the topic and focus of this dissertation, I take the data and its structure as my starting point.

CAS predicts crimes through statistics based on three data sources: BVI (Central Crime Database), GBA (Municipal Administration), and CBS (Demographics from Statistics Netherlands) (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 177–80). The first type of information is socio-economic data from CBS. This data focuses on people’s ages, incomes, and the number of social benefits receivers in an area. The second type of data comes from the *Basisvoorziening Informatie* (basic information provision, BVI), which hosts information that is gathered by the police force itself and focusses on crimes, locations, and known criminals. The third type of data comes from the Municipal Administrations (BAG). The data from this source consists of streets and addresses, which are not used as predictors, but as a structure for the map on which predictions are made. In addition, it often functions as a basis for CBS and BVI data (see Figure 33), since BAGs also consists of basic information about all Dutch citizens such as, for example, social security number, age, and race-ethnic origins.

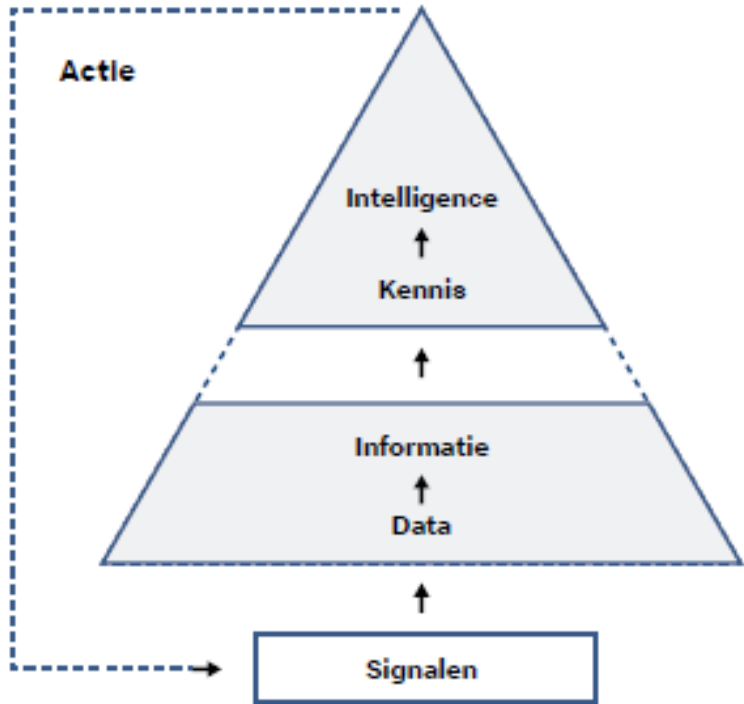


Figure 32: Cycle of information in the CAS data assemblage as shown in (Kop and Klerks 2009, 8). Translation from bottom to top: Signals → Data → Information → Knowledge → Intelligence → Action

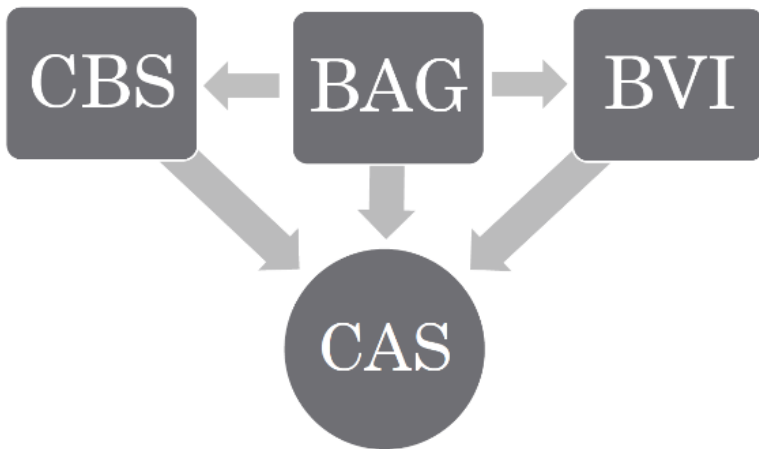


Figure 33: Interconnections between different databases used for CAS

Thus, with its connections to BAG, CBS, and BVI, CAS is embedded in a national and institutional data infrastructure. It should be noted that the CBS, GBA, and BVI databases were not made with CAS in mind, but rather the other way around. CAS repurposes the data collected for statistical and bureaucratic purposes, and is, therefore, tied to the same categorization structures and formats these data have. As I have shown in Chapter 1, CBS statistics were created within a specific national and historical context producing specific definitions and categorizations relating

to the migration background, welfare status etc. of people. Similarly, BVI data emerged within a specific organizational context; an organization with longstanding ideas about the relation between crime and personal characteristics—including race-ethnicity—of people.<sup>76</sup> Both CBS and BVI data are recontextualized in CAS through this process of combination.

The different numbers that constitute the basis of the CAS prediction are called indicators, which in the context of this dissertation can be understood as synonymous with proxies. As explained by Theodore Porter, an indicator

typically cannot measure the very thing of interest, but in its place something whose movements show a consistent relationship to that thing. Since its purpose is merely to indicate as a guide to action, ease of measurement is preferred to meaning or depth. (Porter 2015, 34)

Similarly, in CAS, indicators are an approximation; they are derivative of reality by definition. Since “crime risk” in itself is a construction, a virtuality that has not actualized yet, it cannot be measured directly and, therefore, has to be inferred on the basis of secondary indicators. The best indicators for such an inference are indicators that are themselves measurable and have shown that they predictably relate to the phenomenon of interest. In the choice of indicators, the police holds itself to three standards: (1) indicators have to be *technically compatible* with the data warehouse of CAS and fit the 125 x 125 meter square structure, (2) indicators have to be *tactically relevant* to increase their predictive value and acceptance of results, and (3) the police should be *legally allowed* to use the indicators in question (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 30). While these standards might seem to be logical and suitable rules to adhere to, they also show that the selection of data is based first on availability and compatibility, then on relevance, and finally on legality. Justifiability on other grounds, such as, for example, ethics, equity, or scientific proof, for relations between specific indicators and crime, do not seem to be part of the decision model in the evaluation of CAS in 2016.

An additional point that is often overlooked when reviewing the creation process of a data system is that the data used for the system is seldom explicitly gathered for that purpose. Ideally, when creating a tool, programmers determine which data would be best to have for the predictions or evaluations that they would like to create. If possible, this data is then gathered in order to make ways of measuring and the data format fit the intended goals while simultaneously ensuring that possible biases and errors will be avoided whenever possible. Due to time and budget constraints, this ideal workflow is rarely put into practice, and programmers are often working

---

<sup>76</sup> See the paragraph titled *From “Known Offenders” to “Suspects”* in this chapter for a discussion of this organizational discourse.

with data that already happens to be part of existing bureaucratic practice, often collected with entirely different intentions. Such *exhaust data* are “inherently produced by a device or system, but are usually a byproduct of the main function rather than the primary output” (Kitchin 2014b, 7). In the case of the police, for example, BVI is a database in which information from a collection of different national and international databases<sup>77</sup> is made available for police officers and analysts to be searched via a system that is referred to as BVI-IB (Andersson Elffers Felix 2014, 10). Searches can be done via a variety of different search terms related to, for example, license plates, ID numbers, names, etc. Information is added to these databases on the basis of historical work processes and what officers in the past found sufficiently relevant to enter into these databases. Both the information available, as well as the structure in which it is available, are not necessarily ideal for the purposes of prediction but rather the best compromise in terms of cost and accessibility. In the following sections, I will discuss the different data sources and how they are implemented in CAS.

The first available source of data for the CAS systems is GBA, a municipal database in which all inhabitants of a municipality are registered with their addresses and zip codes. This information is used to create a map on which risk scores are visualized. This map is divided into blocks of 125 x 125 meters, a size which was determined through a process of trial and error by halving squares each time, starting from 1000 x 1000 meters. 125 x 125 meters proved to be a good compromise between precision and practicality. Too big a square would be unworkable for patrolling officers; too small a square would have a low hit rate and render it useless in the operationalization of crime predictions (Willems 2017). After plotting the squares on a map, all ‘empty’ squares, such as squares without houses or companies, and squares consisting of only water, forest, or farmland, are removed. Then of each square, it is determined to which zip code it belongs and which addresses fall within its borders.

Next, historical crime data from the BVI database, consisting of addresses, locations, and time/date of recent incidents of a specific type, are added. For accurate predictions, it is necessary to have a relatively precise time of previous incidents. For pickpocketing and mugging, the time is usually quite precise since victims are usually aware of when these take place. For burglary, however, it is often much harder to determine since victims are often not home or not awake during the incident. Since BVI only works with exact times a rather than with a time

---

<sup>77</sup> While it is not the main subject of this chapter, it should be noted that CAS, or even ILP in the Netherlands, does not exist in isolation from developments in other countries. For example, the city of Berlin (Germany) has implemented a system that is similar to CAS (see A. Meijer, Lorenz, and Wessels 2021). Furthermore, given the ways in which national and international databases are already connected, a more international approach to predictive policing might emerge in the future.

window, the time exactly in the middle of a time window is registered in the database (Willems 2017). In practice, this usually is the time between leaving and returning home, or between falling asleep and waking up. It is unclear whether or not this affects the outcomes of the predictions. Another limitation of CAS is the number of incidents of a given type that happen within a delimited timeframe. Incidents that do not happen very often, such as murders, or incidents that are not always reported, such as rape or the sale of fake drugs, cannot be predicted accurately within a spatiotemporal system. CAS is, therefore, only suitable in targeting relatively minor criminal offenses that happen relatively often in the broader scope of police tasks. A similar focus can be found in BVI and, as I will discuss in the next section, the chosen demographic data from CBS.

Because CAS is operating in an assemblage that emphasizes locations and times, it was specifically made for the prediction of crimes that can be tied to a particular place and time, such as burglary, pickpocketing, mugging, dealing drugs, etc. It is not made to predict or identify incidents that are less explicitly tied to locations, such as most types of white collar crime, like fraud, extortion, etc. While these different types of crimes are also increasingly investigated and fought through digital means (see for example Olsthoorn 2016), these are not the scope of this chapter. It should, however, be noted that, while the popular phrase “predictive policing” could, in theory, signify the computationally aided prediction of any type of crime, in practice, it usually refers to practices in which a computer system is used to manage police officers on patrol in fighting relatively minor crimes that happen relatively often. As a result, predictive policing assemblages are usually aimed at a specific socially, and economically disadvantaged demographic that happens to overlap with the people that commit frequent, minor criminal acts the data system can quantify easily in terms of location and time.

### **Demographic Indicators**

In order to make predictions about locations, CAS makes use of open demographic data provided by CBS (see Figure 34). After determining the zip code of which a square is part, aggregated information regarding that zip code area is incorporated in the risk score of a square. CBS indicators 1, 2, 3, and 4 refer to the number of inhabitants, men, women, and households in an area. Indicators 5, 7, 8, 9, 10, and 11 refer to the composition of the households, the average number of people in a household, any children living at that address, one- or two-parent household. Indicators 11 and 12 refer to the property values in an area and how many properties are currently empty. The next three indicators (14, 15, and 16) are about the number of people that receive an income, including social benefits, and the total income in an area.

Given the size of the squares, the system functions most optimally in urban areas; in rural areas, the number of people and households might not be statistically valid. In addition, a predisposition can be seen towards more normative shapes of families since the system distinguishes between

single-parent and two-parent households and economic factors. This preoccupation of CAS with household composition gives the impression that the police consider both criminals and/or victims more likely to be a part of what is often described as a “broken home.” Because of the indicator regarding the recipients of social benefits, it seems as if people with lower socio-economic status have a higher likelihood of criminal behaviour or victimhood, as the other socio-economic end of the spectrum is not part of the system (i.e., there is no indicator mentioning the number of millionaires).

Nr.	Explanation
1	Number of inhabitants in the zip code area
2	Number of men in the zip code area
3	Number of women in the zip code area
4	Number of households in the zip code area
5	Average size of households in the zip code area
6	Number of non-Western “allochtonen” in zip code area ( <b>removed before 2017</b> )
7	Number of one-person households in the zip code area
8	Number of one-parent households in the zip code area
9	Number of multiple person households without children in the zip code area
10	Number of two-parent households in the zip code area
11	The average age in the zip code area
12	The housing stock in the zip code area
13	Average property value in the zip code area
14	Number of income recipients
15	Number of social benefits recipients in the zip code area
16	Fiscal monthly income in the zip code area

Figure 34: Indicators used in CAS, acquired from CBS.

The final indicator that should be mentioned is indicator number 6 (see Figure 34), which consists of the number of what was then called *niet-Westerse allochtonen*, or people with a non-Western migration background. The police removed this indicator from CAS in 2017. For this part of CAS, the system was tied to the already existing CBS data infrastructure. This means that most statistics were created within the “allochtoon”/“autochtoon” dichotomy. What is striking in CAS 2015 is not only the presence of an indicator for the number of “niet-Westerse allochtonen,” but also the absence of an indicator for “Westerse allochtonen” or “autochtonen.” In CAS 2017, the indicator regarding “niet-Westerse allochtonen” was removed from the systems. While no official reason was provided for this change, literature on the relationship between migration background and crime does make several suggestions.

First, race-ethnic origins could be unsuitable as a predictor for crime. A 2014 study on the criminal behavior of youth aged 12-18 in the Dutch city of Rotterdam could not find a relationship



between race-ethnic origins and criminal behavior but instead found that the specific neighborhoods where people lived were a determining factor (see Driessen, Duursma, and Broekhuizen 2014). Second, the race-ethnic origins could not be of any predictive value since possible suspects are still implicitly woven in the system through the other indicators, making it unnecessary to have race-ethnicity in the list of indicators explicitly. CAS 2017 could still be race-ethnically biased while no longer explicitly taking into account the race-ethnic origins of people living in a neighbourhood, implicitly, through the other demographic indicators that relate to income and family composition, and the location-based approach. After 2016, when more specific indicators regarding household composition, income, and known offenders were added, indicators regarding race-ethnic origins might have been rendered obsolete, either because race-ethnicity does not predict crime or because race-ethnicity is entwined to a large extent with other, better predictors also present in CAS—i.e., these other indicators could, when taken together, be a proxy for race-ethnicity. In this sense, the indicator regarding origins in CAS 2015, at the very least, indicates the ways of thinking of police officers pre-2015, something I will address in more depth in the next paragraph on BVI and its relation to the choices made by police officers in the street.

Another implication of the understanding of crime as a phenomenon through an algorithmically calculated combination of different indicators in CAS is that it works in an inherently intersectional way. The concept of intersectionality points at the way in which various identity characteristics overlap and, as a result, shape specific experiences of oppression (see Crenshaw 1990). In the case of policing in general, and CAS in particular, race-ethnicity and class are used in conjunction with each other and, therefore, create and worsen the oppression of people that are disadvantaged both in terms of race-ethnicity *and* socio-economic status. This “transcoded form of intersectionality” (Cheney-Lippold 2017, 78) plays an important role in the construction of “crime risk” as an indicator, and as I will show in the next section, is in a constant feedback loop with less technologically advanced biases in the work practice of the Dutch police.

### **From “Known Offenders” to “Suspects”**

All indicators extracted from the BVI database refer to data about previously convicted criminals or so-called “known offenders.” For each square in CAS, first, the distance to the home of the closest “known offender” is calculated (see Figure 35, indicator 1). The remaining two indicators refer to the number of known offenders active within the past six months that live within 500 and 1000 meters of a square respectively. In the indicator list of CAS 2017, the term “known offenders” has been replaced with the term “suspects” for categories that still bear the same definition. Although the label still refers to precisely the same category, the word “suspect” has a distinctly different meaning than the label “known offender.” Legally speaking, the word “suspect” is reserved for people of whom “a reasonable suspicion of guilt to a criminal offense is assumed through facts and circumstances” (CBS 2018a, translation by the author). This

classification is, therefore, not valid before a crime has been committed. One could even question whether or not the use of these indicators is in conflict with the principle of letting someone start over with a clean slate after punishment and rehabilitation.

**Nr. Explanation**

1	Distance in km of the address of the closest known offender (suspect) of an incident who has been active in the past six months
2	Number of known offenders (suspects) of an incident who have been active in the past six months and that live within a 500-meter radius
3	Number of known offenders (suspects) of an incident who have been active in the past six months and that live within a 1000-meter radius

*Figure 35: BVI indicators used for each CAS square.*

By analyzing data mentioned in the BVI-section of CAS, we can see how the police understand criminal behavior and which contextual aspects are regarded as important or influential to this criminal behaviour. By discussing recent anthropological and ethnographic work in the section about ethnic profiling, I will argue that the indicators used to predict crime in CAS are not new discoveries but rather datafied continuities and combinations of indicators that police officers have used in their work on the street for a long time.

**Workflow**

The predictions made by CAS are based on the types of data that have been discussed in the previous paragraphs. These indicators, maps, and time frames are first transformed into a visualized prediction by the CAS algorithm. In a second step, they are interpreted and ‘enriched,’ a process of adding information, by data officers. In this section, I will discuss these two steps and explain how they are embedded in the CAS assemblage.

CAS visualizes results in a standardized way for all police departments throughout The Netherlands. The system constructs so-called heat maps—maps illustrating which places have a higher risk for high-impact crimes. These heat maps show blocks that are left blank when the risk is either low or nonexistent; similar to the case of Leefbaarometer discussed in the previous chapter, the colour codes become darker when the calculated risk of a crime increases. Different colours correspond to different crimes (see Figure 36). Only the top three percent of high-risk areas will be colour-coded to relate the number of coloured fields to the capacities of a police force. Tagging more than three percent of coloured areas would be unworkable for the average police district. In addition to the heatmap, CAS provides a simple line-graph where the x-ax shows time (of a single day or week) and the y-ax shows the likelihood of a high-impact crime happening. The graph that shows time and risk is relatively easy to read, but this might come at a cost. Because of the aforementioned problem to determine the time and date of a burglary, the police use an average time in their report. When a burglary victim leaves their home at 8 A.M.

and comes home at 6 P.M., the crime report will state that the burglary happened at 1 P.M. (Willems 2017). Eventually, the graph in CAS may not show a reliable prediction for when the risk is high for a crime such as burglary, but more when the risk is higher of “people being halfway through their workday.” Because CAS functions as a black box, the information officers using CAS, who do not have the option to look up the data used in the construction of this graph, can therefore not reflect on which correlations cause CAS to make particular predictions, even when they would consider this relevant.



Figure 36: Picture of the interface of CAS. Colour code for risk factors: red = burglary; green = street robbery; blue = disturbance by youth; pink = bicycle/scooter theft.

The second visualization within CAS is a grid map (see Figure 36), which is often used in other predictive policing systems. As stated before, the squares with the highest opacity represent the areas with the highest risk for a particular crime to take place, less opaque squares have medium risk, and nearly transparent squares have a low but mentionable risk. The only specifics or filters a user of CAS can add within the grid map visualization is a type of crime. This map will then only show the predicted amount of risk for that selected type of crime. A user cannot, for example, select a square or a crime and receive more in-depth information about the factors which are most important in determining the prediction. In other words, the user cannot see if a risk is predicted on the basis of a stable variable or because of a more changeable or unstable variable, such as the presence of a known offender. Especially in the early stages of CAS, often the same three percent of squares would light up, leaving the impression that some combination of undeterminable factors was causing the risk always to be high for specific areas.

The role of the data officer is to use the visualizations in CAS as a starting point, and from there, to try to explore and retrieve the background of the risks in certain areas (Doeleman et al. 2019). Data officers are able to use other databases or sources from within police departments. At the end of their search, they come up with possible explanations for specific risks. In a way, they

enrich the results presented by CAS because they do not only use the results from CAS or information from the BVI but also use their contextual knowledge about the area and their expertise on crime to look for possible causal effects. In addition, they can add unpredictable events that are not part of CAS, such as football matches of the city's home team or music festivals, and determine autonomously to what extent this influences the prediction. It should be added that the function of visualizing large amounts of data is to make information more accessible and readable for people. This is also the function of the graph and the grid map within CAS. Processing and calculating all the variables and data as quickly as CAS can, would, after all, be impossible for a police officer. However, data visualizations are inherently simplified, and some elements are left out or become obscure (Drucker 2011; Kennedy and Hill 2018). Trying to understand something as complex as crime by consulting a simple visualization could be challenging and is prone to a loss of nuance and misinterpretation (for examples from the Los Angeles Police Department, see Brayne 2021, 56–73). The NPC argues that giving data officers little extra information with the visualization and specifics of the predictions made by CAS works more efficiently. The idea is that since data officers have to rely on their own expertise/experience rather than being closely guided by the results of CAS, they are stimulated to research identified risk areas in a more open-minded way. Data officers, therefore, do not see why a square is coloured, and are forced to investigate this themselves, for example, by checking if crimes have occurred in that area recently or whether something else has been going on (Doeleman et al. 2019).

One could conclude through this analysis that the CAS interface stimulates users to search for additional information from different databases and sources to enrich the predictions and maps made by CAS. The visualization of results within CAS is kept simple, and the numbers of options given to users to explore these visualizations are small. In other words, the Dutch police have consciously chosen to keep this system relatively uncomplicated for users, not providing much additional information on high-risk areas. Therefore, we should not merely consider the ideal work process of the data officer but always consider the emerging risks of police employees skipping the time-consuming process of data enrichment. The very availability of the system, combined with affordances and limitations that keep its users in the dark on purpose, create the temptation of uncritical use of predictions. It should be noted that the larger context of the current discussions on issues of work pressure (see Torre et al. 2011, 29–33), the shortage of personnel within the Dutch National Police (see NOS 2017; R. Meijer 2017), and the interaction of these developments with datafied practices is beyond the scope of this investigation. However, recent research indicates that people working in algorithm-assisted jobs, when in doubt, and, more importantly, under time pressure, tend to trust an algorithm more than their own expertise (see for example, Eubanks 2018 on the “Alegheny algorithm”). Police officers, policymakers, and researchers should, therefore, be advised to continue to pay attention to how various algorithmically aided work processes are translated from policy into daily practices. In

the next section, I will discuss research on how race-ethnicity is operationalized in the daily work practices of Dutch police officers.

## **CAS in the Context of Ethnic Profiling in the Netherlands**

The relevance of race-ethnicity in Dutch policing practices has been a controversial topic for a rather long time. Reports about possible discriminatory behaviour of police officers can be found from at least as early as 1978 (see Esmeyjer and Luning 1978). However, until today evaluation reports still cannot give any conclusive evidence for or against the existence of discriminatory practices. One of the limiting factors in the possibility for investigations into racial profiling in the context of Dutch policing is the lack of a structural effort to gather information about reasons for police to stop, search, and arrest a person, combined with data about that person's appearance or migration background. Furthermore, since Dutch researchers generally feel an uneasiness concerning the concept of race <sup>78</sup> and, therefore, tend to avoid it, the use of "race" and skin colour in profiling practices is often defined in terms of "ethnicity", making it easier for organizations working with this term to deny working with racial categories. As a result, studies about discriminatory policing practices in the Netherlands usually frame this in terms of "ethnic profiling" (see for example Amnesty International 2013, 5). To understand the organizational and cultural context in which CAS has been implemented, we should, therefore, take into account the available recent research on the topic of "ethnic profiling". In this section I will discuss recent anthropological and ethnographic studies from Sinan Çankaya and Paul Mutsaers, as well as police evaluation reports, and reports from NGO's such as Amnesty International.

In his extensive ethnographic fieldwork, cultural anthropologist Sinan Çankaya joined police officers on patrol and observed their decision-making process (Çankaya 2012; 2015). He found out that what he calls "street-level bureaucrats" use three physical categorizations to determine whether or not to stop and interrogate someone (see Figure 37). The characteristics of what Çankaya calls the "biological body" and the "disobedient body" can be seen as real-life parallels of the demographic information of CBS and the "known offenders" data from the BVI. Police officers match this information with the time and location to determine whether or not there is a mismatch and a need for action (Çankaya 2012, 74–76). In addition, Çankaya found that race-ethnicity plays an important role both in decision making during patrols and in the self-identification of the police force as white and autochtoon (see Çankaya 2015; Gowricharn and Çankaya 2017). When we take the work practice and organizational culture of the police, as described by Çankaya, into account, it becomes clear that CAS closely resembles the datafied

---

<sup>78</sup> Further discussion on this uneasiness can be found in Chapter 1 in the paragraph about racialization.

extension of the physical practice of profiling, which includes personal characteristics that relate to race-ethnicity.

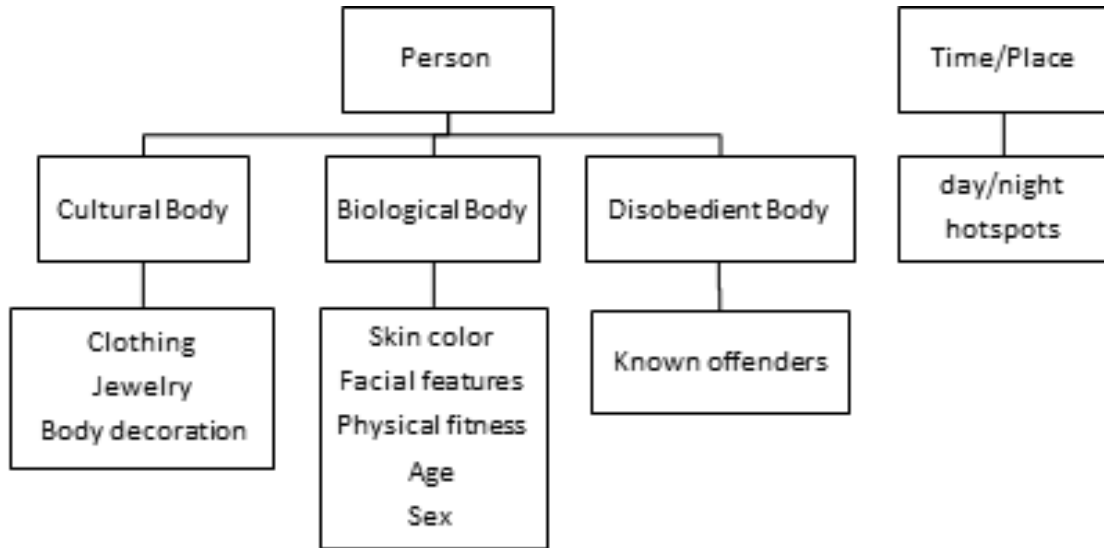


Figure 37: Part of the decision-making process of Dutch police officers (Çankaya 2012, p. 78)

The findings of Sinan Çankaya are corroborated by several other studies of scholars, NGO’s and the police themselves. In a report on the risks of “proactive” policing, Amnesty International (2013) advises the Dutch government against proactive policing policies, which according to them, increase the likelihood of ethnic profiling. Van der Leun and van der Woude (2011, 444–45) come to a similar conclusion in their investigation into the increased interest of the police in preventing crime. The increased possibilities for police officers to use their discretionary powers could come at a cost. Discretionary powers allow officers to make their own decisions without any necessity for substantiation, meaning they do not need to follow any protocol or rules in their decisions to stop and search an individual. In theory, this increases the risk of personal, and by extension, organizational biases against stigmatized minority groups (Eijkman 2010; van der Woude and van der Leun 2013). In follow-up research, empirical data showed no evidence of ethnic profiling (van der Leun et al. 2014). The researchers do, however, recognize that, due to a large number of migrants in two of their three neighborhood measurements, “migration background is not a distinguishing characteristic” (van der Leun et al. 2014, 41). This might indicate that police officers do not single out people individually, but rather target people with a migration background as a group. In a 2021 report, however, after comparing data about self-reported criminal behaviour and statistical information about the ethnicity of crime suspects, sociologists Willemijn Bezemer and Arjen Leerkes (2021) concluded that—after controlling for criminal behaviour—youth with a migration background was between 2 and 3 times more likely to be a suspect in a criminal investigation. To clarify, this means that for the same behaviour, youth with a Dutch background is between 2 and 3 times less likely to be suspected of a crime.

Bezemer and Leerkes clearly state that this is most likely caused by ethnic discrimination, but are unable, due to methodological limitations, to comment on whether police officers or concerned members of the public calling in disturbances are the main source for this discrepancy (Bezemer and Leerkes 2021, 75).

Furthermore, while working with the police in The Netherlands, anthropologist Paul Mutsaers (2015) found that migrant groups were often deemed “target groups” and identified that they were often treated differently from people with a Dutch background in similar situations. For example, during routine checks in an “average” Dutch neighborhood where 6% of the population had a migration background, people who forgot their drivers' license were allowed to call their spouses to bring it over and got away with a warning. In a lower-class neighbourhood in which about half of the people had a migration background, a zero-tolerance policy was applied, and even people with broken bicycle lighting were fined (Mutsaers 2015, 43). The definition of migrants as a “target group,” combined with a less lenient approach in terms of punishment, are ways in which people with a migration background show up more in crime statistics and conversely end up in computer systems as targets, “suspects,” and “high risk.”

Popular and political discourses are partly complicit in shaping crime statistics. Referencing Fabienne Brion (2004; 2005), van der Leun et al. explain that crime databases are not suited for the investigation of crime statistics, since they are the “result and product of official discourse” (van der Leun et al. 2010, 111). Therefore, the share of those statistics that should be ascribed to migrants too, is not only the result of the supposed criminality of particular groups in society, but just as well reflects the significance of race-ethnicity in how police understand crime. Because the police organization, individual police officers, and CAS are largely aligned in the basis for their decision-making processes, to speak of the *risks* of creating feedback loops that place marginalized people at a disadvantage is rather euphemistic. The organizational focus on economically disadvantaged people and the practice of race-ethnic profiling *combined with* datafied risk management can quite literally be understood as “cumulative disadvantage” (Gandy 2016). Cumulative disadvantage creates and reinforces differences in the quality of life of different groups. Discrimination in this sense automatically creates inequality which builds upon itself, hence creating cumulative disadvantage (Gandy 2016, 55).

## **Predictive Policing and Technologically Mediated Racialization**

As I have shown in the previous sections, race-ethnicity has played an—often tacit—role in the ways in which police personnel was managed and police officers make their judgments on the streets. In addition, at the inception of CAS, race-ethnicity was explicitly made part of the indicators that were used to predict crime. However, while the indicator “percentage of non-Western allochtonen” was removed from the CAS system in 2017, the data system is still in interaction with the other parts of the CAS assemblage. The removal of one indicator, therefore,

did not mean that the process of racialization mediated by the CAS assemblage was eliminated then and there. Unfortunately, official information about how policing practices that include CAS relate to race-ethnicity and whether or not these practices still feature forms of technologically mediated racialization is absent. However, despite the absence of structured information about ethnic profiling in the work practice of Dutch police officers we can conclude that, to speak in algorithmic terms, the combination of CAS and the organizational culture of the Dutch police are 1) at risk of racial and ethnic profiling practices, and 2) lack proper institutional structures to prevent racial and ethnic profiling.

On an instrumental level (i.e. approaching a system in terms of its purpose and efficiency), CAS was explicitly connecting crime to race-ethnicity in the period before 2017. It did so by employing one category in particular, “percentage of non-Western allochtonen,” as an indicator in its assessment of the risk of burglary, robbery, and other “high impact crimes” committed in areas of a city. In doing so, CAS used the race-ethnic category as the digital variant of a stereotype, increasing the chance of people with a non-Western migration background of being subjected to surveillance while decreasing the chance of people with a Western or Dutch migration background to be subjected to a similar treatment. When looking merely at CAS as an individual data system, the removal of race-ethnic categorizations in the iterations of CAS after 2017 seems to reduce the risk of a datafied form of ethnic profiling. However, by considering CAS in relation to the available research on the prevalence of ethnic profiling by Dutch police officers, it becomes clear that race-ethnicity as a factor can easily seep into the calculations of a system like CAS.

On an epistemological level (i.e. considering a system as a knowledge producing agent with a particular perspective), I have shown that CAS does not exist apart from cultural and societal norms. As Lisa Gitelman emphasizes in the title of her edited volume “Raw data is an oxymoron” (Gitelman 2013), all data is situated in a particular history, context, and institutional form, which together make up what is seen as relevant, irrelevant, subjective, or objective. The misconception here is that data-driven systems are based on data that includes everything and/or is exhaustive. As Rob Kitchin (2014a, 4) states clearly:

Big Data may seek to be exhaustive, capturing a whole domain and providing full resolution, it is both a representation and a sample, shaped by the technology and platform used, the data ontology employed and the regulatory environment. (Kitchin 2014a, 4)

This is also the case with the data selection in CAS. Choices of certain data sets are made on the basis of technical compatibility, tactical relevance, and legal grounds. This creates a significant reduction in the number of available indicators that were severely limited in their coverage of real-world phenomena to begin with. While the chosen indicators might be relatively good in relation to what is available, the choice is, nevertheless, always shaped by existing categorizations, bureaucratic practices, and epistemic culture.



In addition, one could add that the selection criteria still lack any ethical consideration and consideration of what happens once data are entered into a predictive system. With reference to legal scholar Lior Jacob Strahilevic (2008a; 2008b) and statistical epidemiologist Sander Greenland (2002), communication scholar Oscar H. Gandy (2016, 62–65) argues that some indicators, such as race or ethnicity, could best not be used in predictive systems, not because they are bad predictors, but because it would be unethical to use non-causal factors. While race-ethnic indicators were removed from CAS in 2017, it could be argued that this line of reasoning also applies to other indicators, such as, for example, the number of welfare recipients in an area. Not only would the intersectional nature of algorithmic calculations turn other indicators into proxies for race-ethnicity, they could and would themselves become new human characteristics available for “purposes of racial signification” (Omi and Winant 2015). This shows one of the fundamental criticisms of predictive policing and other forms of “statistical discrimination” (Gandy 2016), as the discussion should not be about this or that data point, but rather about what the process of predictive policing does to those data. In this understanding, as explained by the media scholar Joshua R. Scannell, it is impossible to remove race from policing, because “race is an originary policing technology, just as policing is a bedrock racializing technology” (Scannell 2019, 108). It is the process of technologically mediated racialization that can, and will, turn any human characteristic that it can process into a data point that is available for racial signification.

A similar argument against predictive policing is made by Peter Mantello about what he calls “the precrime assemblage”:

Rather than creating more equitable and fair law enforcement practices and judicial processes, the precrime assemblage judges persons not as individuals but rather as numerical signifiers—placing them in actuarial categories based on what others have done in the past. (Mantello 2016, 9)

The use of target groups or “actuarial categories” and population characteristics or “numerical signifiers” in data-driven systems hide “racial, religious, and often socio-economic disparities behind a veneer of scientific impartiality” (Mantello 2016, 8). This means that a system without data about population characteristics is not automatically neutral or fair for people with different migration backgrounds. As Lum and Isaac show in their analysis of the PredPol algorithm, a system that only uses the three data points crime type, location, and time, still produces a racially biased prediction (Lum and Isaac 2016). Similarly, CAS is built upon an existing bureaucratic infrastructure of NPC, which includes the cultural norms of the organization and its officers. As a result, CAS will only be as equal towards all Dutch inhabitants as the data and infrastructure it is built on and as the work practice of the police officers that are managed by it.

Unfortunately, currently available research, especially on the practice of police officers is not undecidedly positive (see Çankaya 2015; Mutsaers 2015). Unlike the race-ethnicity of convicted

criminals, quantitative information about the attitudes of police officers towards people with a non-Dutch migration background is not structurally measured. Despite numerous reports by human rights organizations (see Amnesty International 2013; Controle Alt Delete 2019) and the results of qualitative studies about this topic (Çankaya 2012; 2015; Mutsaers 2015), information on the scope and ubiquity of ethnic profiling done by the Dutch police, is, still, inconclusive (see van der Leun et al. 2014; Molina Espeleta et al. 2021). A recent noteworthy effort to have all police officers file reports on their interactions with citizens in a dedicated app failed because registrations are not mandatory, and most police officers do not see the urgency or feel the need to register such information (see Molenaar et al. 2020). The current practices of the Dutch police with regards to race-ethnic matters are, therefore, rather opaque, and the lacuna in knowledge makes it hard to hold this organization accountable for its actions.

Furthermore, on an ontological level (i.e. considering a system in terms of its world and life creating capabilities), the lack of information risks the creation of unequal treatment. First, people committing the same crimes could end up having very different chances of getting caught, which would further feed the system with biased information. Second, when crime becomes racialized, particular behaviour becomes suspicious for racial reasons rather than criminal ones. A well-known historical example of this is the concept of “mugging,” which became reserved for robberies committed by black youth in London in the 1980s (Hall et al. 1978). Stuart Hall and his colleagues showed how the prioritization of “muggings” by the London police was just as much guided by the skin colour of the delinquents as by the severity of the crimes (Hall et al. 1978, 325). It should be understood that this does not merely lead to unequal treatment of criminals but, more importantly, to the attachment of racial meaning to particular crimes. If only robberies committed by black people are considered muggings, and mugging would be quantified by a predictive policing system, we would end up with a system that is racially biased without any explicit reference to race. Due to the lack of information available about how CAS works and interacts with police officers in the street, there is currently no way of knowing whether CAS features datafied variants of racial stereotypes and/or datafied variants of behavioural concepts that are invested with racial meaning.

The lack of accountability of the Dutch police in matters that relate to race-ethnicity creates a situation in which race-ethnicity is a tacit presence woven into the fabric of the predictive policing assemblage. By refraining from making this possibility less likely, the police acts as if the risk of producing race-ethnic bias is one worth taking. Since predictive systems make their predictions through the operationalization of the apparent unequal distribution of criminal behaviour in society, risking this possibility is especially unfair. Instead of aiming to solve the unequal distribution of crime, predictive policing assemblages are aimed at exploiting this inequality. As explained by digital media scholar Wendy Hui Kyong Chun, by considering the unequal spatial

and social distribution of people a naturally occurring form of homophily, digital systems introduce

normativity within a supposedly nonnormative system by presuming that consensus stems from similarity; and it makes segregation the default. In valorizing “voluntary” actions, it erases historical contingencies, institutional discrimination, and economic realities. (Chun 2021, 92)

In short, in order to function, predictive policing assemblages normalize and need inequality. In that sense, predictive policing practices are at odds with any initiative that tries to decrease inequality. After all, if an initiative, such as welfare redistribution, results in an equal distribution of criminal behaviour, risk profiles become random and, hence, less effective. Arguments such as these clearly show how questions with regard to “efficiency” cannot be separated from questions with regard to justice and equality. In other words an increase in “efficiency” most likely points at increased inequality, rather than at a “better” algorithmic system. In a society that claims to strive for equal opportunities, the “efficiency” of a datafied policing assemblage is, therefore, something that should be avoided, since it is most likely entangled with an increasingly unequal society.

## Conclusion

In this chapter, I have discussed yet another case in which race-ethnically categorized data made available by CBS were appropriated by a governmental institution. I have shown that, while these data have since been removed from CAS, the possibility for technologically mediated racialization to occur, remains intact. By looking at CAS as part of a predictive policing assemblage that includes the practices of patrolling officers, we can see that CAS is at the mercy of the ethical and racial awareness of the people that work with it. Given the anthropological and ethnographic evidence about the organizational culture of the Dutch police (Çankaya 2015; Mutsaers 2015), the CAS assemblage should explicitly *not* get the benefit of the doubt when it comes to ethnic profiling. Furthermore, as I have argued, the specific process through which predictive policing assemblages produce meaning about the human characteristics they quantify, is in itself a racializing process. Systems such as CAS are not only “using” race-ethnicity, or “reporting on” race-ethnicity; they create specific configurations of race-ethnicity that are intimately tied to data technologies in a specific cultural context.

To counteract, or prevent, detrimental effects of predictive policing technologies for marginalized groups, we need more thorough processes of accountability pertaining to policing in general, and datafied policing in particular. The most striking part in the case of CAS is not so much the *available* information about it, but rather what is *not available*. Even if we consider CAS without reference to possible ethnic profiling, scientific evidence for the usability, accuracy, and efficiency of predictive policing seems rather hard to come by, not in the least because of some

rather tricky methodological issues that still need to be addressed. When attempting to prevent crime from happening, it is difficult to know if one is effective. After all, after a quiet night of patrolling without any burglaries taking place, no one knows if the prediction of burglary was right or wrong—if crime is prevented, nothing happens; if the prediction is wrong, nothing happens. On the contrary, if police officers do catch a burglar, it is hard to know if the predictive system was right or lucky. It is, therefore, not surprising that a review of known empirical studies of predictive policing systems published before April 2017 did not find any evidence for or against the efficiency of predictive policing and did not find any evidence for or against discriminatory tendencies of such systems (see A. Meijer and Wessels 2019). The 2017 evaluation report on predictive policing practice with CAS did come to similar conclusions as the researchers note that they were unable to find any indications that predictive policing leads to decreasing crime or a decreasing rise in crime (see Mali, Bronkhorst-Giesen, and Den Hengst 2017, 41). Despite these conclusions, CAS is still in use at the time of writing i.e. at the end of 2021.

What does stand out in the 2017 evaluation of CAS is that the NPC and even the researchers seem to be driven by techno-positive and determinist visions. For example, in their discussion of the accuracy of CAS, researchers state that the hit rate of 13,5 percent is 4,5 times better than a random distribution (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 40). While this might seem a good number, especially for a recently developed system, this is not a valuable measurement. It is not the comparison with a random distribution that is interesting, but rather the comparison with more traditional ways of working. Here, I assume that before the introduction of CAS, the police were not just patrolling randomly, but rather used their intuition, and possibly less technologically advanced forms of statistical planning. Furthermore, the writers of this report state, on the one hand, that they could not find any evidence that predictive policing leads to diminishing crime rates, and, on the other, that they do not question *if* predictive policing should be introduced, but rather *how* the police should do this (Mali, Bronkhorst-Giesen, and Den Hengst 2017). Meanwhile, the conclusions of this report are filled with words like “potential,” “opportunities,” and “enthusiasm,” where we would want to read “evidence” and “proof” and “results.” The report states literally that all performance indicators that were measured, namely the number of burglaries, the total number of reported crimes, and the number of mutations added to the system by police officers, did not show any effect that could be attributed to CAS (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 41) It is, therefore, somewhat surprising that while the pilot of the new system concludes that “the subjective perception (...) seems to deviate from the more objective findings” (Mali, Bronkhorst-Giesen, and Den Hengst 2017, 41, translation by the author), CAS was rolled out nationally only a few months after the publication of this report without major changes to its design. As such, CAS is functioning as an assumed self-fulfilling prophecy that, as of yet, still needs to prove itself.

Thus, convincing evidence for the efficiency and usability of CAS is still lacking. Moreover, policy goals, benchmarks, or performance indicators are also lacking.<sup>79</sup> This makes it very hard to prove or disprove that CAS is a valuable addition to the working practice of the Dutch police, which in turn also leads to questions about the legitimacy of the decision to implement the system nationally (see Rienks and Schuilenburg 2020, 45–46). Combined with the lack of quantitative, statistical, and empirical knowledge about the prevalence of ethnic profiling, the lack of accountability about CAS creates a situation in which there is no way to know how the police in particular, and Dutch society in general, benefit from the implementation of ILP and CAS. Whether or not doubts about new systems are legitimate should be of great concern to the Dutch police since the very existence of such doubts already delegitimizes their practice and potentially creates mistrust in the police. It is, therefore, of the utmost importance that the police create suitable checks and balances that can guarantee that this powerful organization can be held accountable in case of accidents and misuse. One straightforward way in which the police could improve its accountability would be to have clear ethical and methodological guidelines for the assessment of new policies, technologies, weapons, and associated working practices. In addition, before future testing, it could be decided which results would be considered a success, and could possibly lead to the implementation of new tools and practice, and which results would lead to redesign and more testing. In addition, certain principles, such as, for example, equity, should be part of the design process from the start. After all, technologies equipped with tasks as important as aiding in keeping people safe, equality should not merely be an afterthought, but instead, a guiding principle.

---

<sup>79</sup> This was pointed out by the Dutch MP Beukering-Huijbregts of the D66 party in the Parliament's Lower House and confirmed by the Minister of Justice, Grapperhaus (see Grapperhaus, Dekker, and Broekers-Knol 2020, question 131).

## Chapter 6: Conclusions and Recommendations

In this dissertation, I have not only investigated how race-ethnicity is known and used, but also how it is produced in relation to governmental data technologies in the Netherlands. I proposed to understand the process of what I termed *technologically mediated racialization* through three different approaches: instrumental, epistemological, and ontological. In this concluding chapter, I will first answer the main research questions. In the second part of these conclusions, I will discuss how debates that are currently going on in most societies that are implementing digital technologies in governmental processes, in particular Dutch society, can benefit from the results of this research. The current situation of data technologies used in governmental practices, and especially the routine use of race-ethnic categories in them, calls for tools aiding in the prevention of algorithmically generated racist outcomes, and guidelines for ethical and anti-racist uses of categories in situations where they can benefit marginalized groups.

Systems aimed at benefitting people from marginalized groups is an important topic that I have not touched upon in this dissertation. Removing race-ethnic categories can be counterproductive, as it would hinder applications aimed at providing possibilities for affirmative action. The solution for this apparent paradox can be found in the idea of strategic essentialism, a well-theorized concept in feminist and postcolonial theory, that points at the possibility for allowing temporary epistemic inaccuracy in order to strive for the achievement of an anti-racist and anti-sexist goal. Thinking through this concept in cases where identity characteristics are used in datafied systems might aid in creating situated anti-racist classifications suitable for the specific goals for which they are invented. Strategic essentialism is, therefore, a promising concept not only for research into race-ethnic classification, but also in policy practice, as an aid in pursuing ethical and anti-racist datafied practices. I will show how my analytical process in this investigation of technologically mediated racialization - from instrumental use, through epistemological critique, to ontological deconstruction - can be inverted and used in the planning, evaluation, and design processes of current and future governmental practices that are aided by data technologies. In the final part of this chapter, I will discuss several possible future research directions that can be derived from the results of this investigation.

### Findings

When I wrote the proposal for this research between the summer of 2016 and the spring of 2017, I did not foresee that during the course of this investigation all the labels of classifications that are object of this dissertation would be changed, and the problematic clusters of migration backgrounds formerly termed “*Westers*” and “*niet-Westers*” would be abolished. The reader should, therefore, be aware that while the research questions invite answers that explain a fixed situation, in reality the objects of investigation are moving targets. I will, therefore, try not only to give answers that pertain to the previous period of about 5 years, but also point at the

theoretical findings that will remain relevant and pertinent, regardless of future categories and labels. For reasons of clarity and readability, I will first answer the sub-questions of this investigation before turning to the main research question. In what follows, I will discuss the sub-questions in the order they were posited in the introduction to this dissertation.

## Categorization

1. What are the origins of the Dutch race-ethnic categorization system and how are its categories defined?

At the end of 2021, the categorization system used by the Dutch government in general and CBS in particular was in a state of flux. Both the labeling practices and the clustering practices that had been the standard for the past few decades were being evaluated and replaced (see the reports by WRR, Bovens et al. 2016; 2021; R. Jennissen, Bovenkerk, and Engbersen 2021). I will, therefore, first reflect on the situation right before these changes were made, and then discuss the effects of the changes on the way in which knowledge about migration background is created by CBS.

The main classifications of people that dominated discourses on migration, integration, and many other societal developments in the Netherlands were the categories “autochtoon” (for people with two parents born in the Netherlands) and “allochtoon” (for people who have at least one parent with a place of birth outside of the Netherlands). It should be noted that former colonial territories, such as Surinam, Indonesia, and the Dutch Antilles, count as outside the Netherlands in this system. Due to the increasingly stigmatizing connotations of the word “allochtoon”, it was decided in 2016 to replace the “allochtoon”/“autochtoon” dichotomy with the more neutral description of *Nederlandse achtergrond*, or Dutch background, and *migratieachtergrond*, or migration background. However, as clearly stated on the website of CBS at least until the end of 2021, while the labels were replaced, their “definition did not change” (CBS 2019, translation by author). After all, the underlying human characteristic, the birthplace of parents, was not altered. In the context of datafication, this means that the functionality of the categories remains unchanged as well. As I will discuss further in the section on the third sub-question, this means that third party data systems did not have to change their algorithm, functionalities, or language either. The conceptualization of migration background *as a data point*, and as a human characteristic made available for racial signification, was not affected.

The now abolished subdivision of the classifications “Westerse allochtonen” and “niet-Westerse allochtonen” was officially based on the “socio-economic and cultural similarities” of countries with the Netherlands. However this division does not consistently follow any specific characteristic of selection, such as geography, language, or status as a former colony. As explained in Chapter 3 and to a lesser extent in the introduction to this dissertation, the specific

way in which the clustering of different migration backgrounds happened can be understood as a racializing division. One tell-tale sign of the connection of these classifications with race are the words of the influential sociologist Hilda Verweij-Jonker, who stipulated that her choice of the collection of migration backgrounds in a report was based on a “striking appearance - especially the skin color” and on the “strange” and for Dutch people “incomprehensible” language that people spoke (Verweij-Jonker 1971, 7). This resulted in a cluster of categorizations in which the label “non-Western” was used from the 1990s onwards, for migration backgrounds of groups that were considered problematic in the Netherlands. Generally speaking, the label “Western” was used for migration backgrounds that fit the ideal type of the affluent, racially white, Christian person. The rather arbitrary application of these principles led to, for example, a different classification for the former colonies Indonesia, that was considered “Western”, and Surinam, that was considered non-Western.

The abolishment of the adjectives “Western” and “non-Western” followed a different trajectory than the noun “allochtoon” did. In 2016, they were first replaced by “people with a Western migration background” and “people with a non-Western migration background”. In 2021, the subdivision was abolished entirely, after a report by the Scientific Council for Government Policy (WRR) explained that there is no scientific substantiation for it, even though the pair of concepts was still used often by researchers and civil servants (Bovens et al. 2021, 9). This is no minor verdict and it does raise the question why it was ever implemented.

The implementation of the classifications in various policy directives and digital applications can, at least in part, be explained by the classifications having been the default structure for many different statistics about socio-economic phenomena. This default structure did not only manifest itself in public discourse about topics where researchers and civil servants felt the need to distinguish between different migration backgrounds, but also in the technical infrastructure made available to citizens, researchers, and journalists. While labels and clusters have undergone a transformation in the past several years, the way in which this information is made available is still very similar. In the answer to the next sub-question, I will address this technical infrastructure in more detail.

## Datafication

2. How are race-ethnic categorization systems in the Netherlands datafied by government institutions?

Like most information collected by CBS, race-ethnic information became increasingly datafied. The process of datafication is understood in this dissertation as the transformation of all aspects of life into quantifiable data (Mayer-Schönberger and Cukier 2013). In the case of CBS, the most significant changes in this regard happened during and after the census of 1970, when, for the



first time, counting no longer took place by visiting people in person, but by adding up the registrations of municipalities. This made the process of counting and processing data much easier and cheaper. Like all other information that was gathered from citizens, race-ethnicity became a data point, available to be matched with any socio-economic or cultural variable with which it can be paired.

On an instrumental level, technological developments of the 80s and 90s, made it possible for CBS to also start sharing statistics with third parties by digital means. First, statistics were made available on physical data carriers, such as CD-Roms, and later via the internet. The creation of an API (application programming interface) created the possibility for third parties to create their own services, interfaces, and knowledge producing applications with CBS data. Another affordance of race-ethnic data available through an API, is that it is easier to combine it with other context-specific data. The uses of such data have become quite widespread, as can be seen from the case studies in Chapters 4 and 5, and in the recent scandal concerning SyRI (see van Schendel 2019) and the algorithmic system used by the Dutch Tax and Customs Administration in the “childcare benefits affair” (see Autoriteit Persoonsgegevens 2020).

On an epistemological level, as can be seen from the case-studies in this dissertation, the datafication of race-ethnicity by an important Dutch institution such as CBS creates its own ethical problems, regardless of the ways in which people are categorized. Since CBS is viewed as an objective provider of information, the very availability of information legitimizes its use. Furthermore, the availability of information in a particular structure—whether it be through the “allochtoon”/“autochtoon” dichotomy or in any other way—creates the impression that this information is in itself neutral, independent of the ways in which it is reused. This neutral and objective appearance is reflected in the different applications in which race-ethnic information has been reused in different parts of the government, such as fraud detection, crime prediction, and the calculation of livability. This shows that governmental organizations treat race-ethnicity like any other datapoint made available for use, seemingly without being aware of the sensitive nature of information about migration background, and the risks associated with its use.

## Appropriation

3. How are race-ethnically categorized data appropriated by governmental and non-governmental organizations?

In theory, the availability of race-ethnically sorted open data makes it useful for many different purposes. The scope of this dissertation has been limited to a few specific cases that can be seen as archetypal situations in which data is repurposed. The first type of reuse is by rightwing politicians and media. As I have shown in earlier work with other colleagues, the interface of CBS allows for the customization of data selection and visualizations, a feature which is regularly used

by rightwing politicians to share socio-economic “facts” about the lives of people with a migration background (see Muis et al. 2019). The affordances of the CBS interface allow for a strategic selection of data to make a situation seem more alarming. Furthermore, the connection of race-ethnicity with any other socio-economic phenomenon gives the illusion of a causal relation, where in most situations the only proven fact is a correlation.

However, as I have shown in the case-study of the Allochtonenmeter, the options of the CBS interface are not limited to the static display of this or that table. The available open data allows for the creation of interactive stand alone applications. In the case of the Allochtonenmeter, a relatively simple search algorithm returns the percentage of “non-Western allochtonen” in a postal code area. Both the “Allochtoon-O-meter” once featured on the alt-right news site *Geenstijl*, and the independent Allochtonenmeter communicate their connection with ethno-nationalist and racist discourse through their interface and accompanying explanations. It could be argued that, in a cruel way, the Allochtonenmeters are producing knowledge that is situated to some extent. At the very least, the applications are rather obvious in communicating their political affiliations and/or non-neutral position, and their only claim to objectivity is constructed by relating the information to the CBS infrastructure.

In contrast, the communication of positionality is nowhere to be seen in government applications such as the *Leefbaarometer*. This application with a map of the Netherlands features several elements, such as the prominent government logo, and a professional and neutral-looking interface, that together communicate an impartial and objective view on reality. However, hidden underneath this sheen of objectivity are indicators and an algorithm that shows similar prejudices as the information presented in the Allochtonenmeters. First, while the *Leefbaarometer* comes of as a measurement of a particular objective phenomenon, it is actually a prediction of a situated socio-technical construction that is called “livability”. This “livability” is, amongst many other indicators, related to the migration backgrounds of people that live in an area. A structural penalty is given for the presence of people with a migration background and the only group not represented in the system is the group that does have an impact on the score: people with a Dutch background.

To make matters worse, at the end of 2021, the explanations of the calculations on the website, still use the abolished and stigmatizing words “allochtoon” and “autochtoon”. Since the data on which the system is based, was updated in 2018, that is, after the imposed vocabulary change, the unchanged labels communicate a particularly ignorant attitude towards issues concerning race-ethnicity. This shows how merely changing labels of data in Open Data infrastructure—like the one provided by CBS, which is explicitly designed to create possibilities for the reuse and recontextualization of data—, is rather irrelevant in a datafied context, as this leaves the function and operationalization of particular data intact. In a datafied society, it is the racializing function of a category, and not so much a stigmatizing label, that does the most damage. Therefore, the

choice to replace labels in 2016, and not the definitions of racialized categories, changed nothing in processes of technologically mediated racialization.

The case of the Crime Anticipation System of the Dutch National Police Corps shows how race-ethnicity can play a role in the performance of a data assemblage, even when race-ethnic data is not explicitly embedded in a data system that is part of that assemblage. The interaction between a predictive system and the work practice of police officers can produce racializing outcomes, due to the possible presence of human prejudices that are not accounted for. While there is abundant social, ethnographic, and anthropological evidence that profiling based on race-ethnicity does happen in the work practice of the Dutch police, there are still no systems or policies in place to structurally monitor the performance of the police in this matter. When such practice is combined with a predictive system, the assemblage as a whole is at risk of producing unfair outcomes for stigmatized groups of people in particular.

## The Practices and Consequences of the Datafication of Race-ethnicity

***Main question: What are the practices and consequences of the datafication of race-ethnic categories in the Netherlands?***

Through the infrastructural inversion that encompasses the history of race-ethnic categorization in Chapter 3, combined with the case studies in Chapters 4 and 5, I have shown that technologically mediated racialization is a distributed process that is enacted throughout the Dutch governmental data assemblage as a whole. However, CBS does play a central, infrastructural, role in this process by providing the technical and epistemological infrastructure that affords data and knowledge distribution. In doing so, it lays the ground work for a conception of Dutchness that is intimately tied to the technological infrastructure that mediates it. In the following three paragraphs, I will go through the implications of this technologically mediated racialization through the, by now, familiar three approaches.

Through the instrumental approach, I have shown that the availability of race-ethnic data in the infrastructure of CBS in itself makes that it is conceived as a regular data point like any other. Like assigned sex, date of birth, and place of birth, the migration background of all Dutch people is routinely taken from municipal registrations and processed into many different statistics. These statistics are then made available by CBS for any person or organization that is interested in reading, recombining, and processing this information. As can be seen in the case studies of Chapters 4 and 5, as well as in the news about recent scandals concerning the discrimination of people with a migration background by various Dutch governmental agencies (see Uitermark, Hochstenbach, and van Gent 2017; van Schendel 2019; Autoriteit Persoonsgegevens 2020), the affordances for this reuse of data were actually implemented in various applications. In most situations, it is unclear why the explicit choice was made to include information about migration

background, and what this information has to add to the systems. In these situations, the instrumentalization of race-ethnicity amounts to the increasing racialization and stigmatization of people with a migration background in general, and people with a “non-Western” migration background in particular. Such an instrumentalization of race-ethnicity in algorithmic systems, therefore, inevitably leads to the production of race-ethnically biased results in situations that should otherwise be treated equally.

On an epistemological level, the datafication of statistics based on race-ethnicity creates a situation in which applications built on the central infrastructure are encouraged to use the same categorizations and clustering as the larger infrastructure that it is connected to. Such a construction forces a standardized race-ethnic layer as a frame on a variety of different social phenomena, ranging from crime and livability to initiatives aimed at improving diversity and inclusion. From a usability, data, and infrastructural perspective this practice makes sense. After all, the possibilities for the recombination and reuse of data points is made much easier when all parts in an assemblage use the same technical standards and definitions. From a sociological perspective, however, forcing a standard frame on a variety of different social and cultural problems in a variety of different contexts might be rather ineffective as it structurally limits the scope of knowledge production to parameters often set years in advance, and very often set from the ideologically dominant perspective. As a result, the categorizations used are most suitable for racializing uses; that is, for practices with which the categories already have discursive connections, such as topics relating to integration, policing, and poverty. The Dutch Scientific Council for Government Policy (WRR) acknowledges this problem by stating that replacing the categorizations and clusters “autochtoon”, “allochtoon”, Westers, and niet-Westers” should ideally not be done with one generic system, but rather by “using a multiplicity of clusters and terms” (Bovens et al. 2016, 66, translation by the author) dependent on the social issue at hand. However, as I have shown in Chapter 4 and Chapter 5, this suggestion completely contradicts the ongoing project of the datafication of race-ethnicity that is currently undertaken by various Dutch governmental institutions. Since their process was, until recently, largely driven by an increasing generalization and infrastructuralization of standardized race-ethnic terminology.

In an attempt to thwart the effects of the now racialized race-ethnic clusters of “autochtoon”, “Westerse allochtoon”, and “niet-Westerse allochtoon” the Dutch government decided to abolish the labels, without removing the categories and clusters. As I have shown in Chapter 4 in particular, this abolishment of labels in a datafied context amounts to a form of colonial aphasia—explained as “a loss of access that may verge on active dissociation” (Stoler 2016, 12)—in two distinct ways. A first form of colonial aphasia can be found in the apparent hard break in racializing discourse that is effectuated by a vocabulary change. When such a vocabulary change is not combined with a change in categorization and clustering practice, racializing processes do not disappear, but become more difficult to address. When the “old” idiom becomes a forbidden

discourse, appeals for anti-racist policies that take into account particular racialized human characteristics, such as skin colour or migration background, start to sound rude and become more difficult to distinguish from racist discourse and conduct. While it is true that people should not be forced into race-ethnic categories by the government, we should recognize that making the discourse “not done” also makes it impossible to address ongoing discrimination in a society in which those race-ethnic identities are made real on a daily basis.

A second form of colonial aphasia is caused by the datafication of race-ethnicity, which creates the possibility for the decontextualized use of race-ethnic data. Race becomes just one of many decontextualized datapoints about which correlation trumps causation. The labels given to datapoints at the beginning of the datafied pipeline are irrelevant further down the line of data processing. Not only are the specific datapoints that are used in an algorithm often obfuscated by the interface and visualizations that communicate the outcomes of a process, third parties that appropriate data for a new purpose can easily continue to use abolished vocabulary if they feel like it. Labels, classifications, and clusters should, therefore, not only be considered within their original context, but also in terms of the possible combinations that they can engage in through practices of datafication; the possibility of so-called *context collapse*<sup>80</sup> should not merely be thought of as a possibility, but as a structuring principle and, therefore, as an inevitable aspect of datafication. This means that if a particular category becomes racialized, we should think about solutions that go much further than merely changing the label. In the final paragraph of this conclusion, I will address a few possible directions for future practice and policy concerning the datafication of race-ethnic categories.

Through the ontological approach, I have shown how the datafication of race-ethnicity does not only aid knowledge about different people in the world in general and the Netherlands in particular, it produces the specific configurations in which those people exist, and are given the opportunity to exist. In this light, the configurations of “allochtoon” and “autochtoon” are, therefore, not only social constructs but *sociotechnical data assemblages* that can operate independent of vocabulary. In the data assemblage that currently produces race-ethnic identities in the Netherlands, the CBS data infrastructure plays a vital part. As it is not only the provider of default categorizations and clusters, but also the provider of the technical infrastructure necessary to distribute the default race-ethnic categorizations. This creates a particular condition of possibility, which I termed *technicity of race* in Chapter 1, through which its discourse on race-ethnicity becomes a discourse *network* on race-ethnicity, capable of exchanging information and

---

<sup>80</sup> *Context collapse* is a concept often used in social media research that refers to “how people, information, and norms from one context seep into the bounds of another” (see Davis and Jurgenson 2014).

meanings with third party systems that directly shape the living conditions of people living in the Netherlands.

Through the API of CBS, the labelling, categorizations, and clustering of its information becomes a default way of understanding various social phenomena in different contexts. While data science and statistics—in this case with the help of race-ethnic data—are usually only capable of producing knowledge in terms of correlations, in practice results are treated as causations, resulting in an unequal distribution of life chances of people living in the Netherlands. The resulting inequalities can be seen in the various recent scandals about data systems in Dutch media (see Uitermark, Hochstenbach, and van Gent 2017; van Schendel 2019; Autoriteit Persoonsgegevens 2020), as well as in the examples from Chapters 4 and 5. Due to the networked nature of the data assemblages that produced these inequalities, the responsibility for these instances of technologically mediated racialization should also be understood in a networked manner. The end users of information should be held accountable for their practices, but we should also address the affordances of the CBS infrastructure that, in part, made these practices possible. It is the CBS data infrastructure that makes Dutch citizens not only knowable as race-ethnic human beings *in a particular way*, it makes us exist as race-ethnic human beings *in a particular way*. As such, we can understand contemporary Dutch race-ethnicities in particular, and racial formations in general as *relational ontologies* in which racial formations are always mediated by specific data technologies; each technology produces a racial ontology based on its historical and sociotechnical context *and* technological possibilities.

When we take all three approaches to *technologically mediated racialization* together, we can see that the particular way in which new technologies are able to record and transport data about race-ethnicity or migration background is in no way random, and never a complete redefinition of what race, or race-ethnicity means in a given situation. Instead, new technologies both inherit existing racialized vocabulary, ideas, and infrastructures, and add on to them. In this light, we can see that the Dutch governmental data infrastructure is not merely a neutral information resource, but rather a link between racialized Dutch history and social conventions, and contemporary racialized government practice. As such, the Dutch datafied society seems to be structured by what could be called a Dutch Racial Contract. In his influential work on this concept Charles Mills explains how The Racial Contract,

establishes a racial polity, a racial state, and a racial juridical system, where the status of whites and nonwhites is clearly demarcated, whether by law or custom. And the purpose of this state [is] specifically to maintain and reproduce this racial order, securing the privileges and advantages of the full white citizens and maintaining the subordination of nonwhites. (...) From the inception, then, race is in no way an 'afterthought,' a 'deviation' from ostensibly raceless Western ideals, but rather a central shaping constituent of those ideals. (Mills 1997, 13–14)

While the limited scope of this investigation does not warrant a conclusion about the Dutch state as a whole, it does suggest that race is not an “afterthought” in the construction of datafied bureaucratic technologies. When my findings are combined with existing recent works about Dutch discourses on race and identity (see, for example, Essed and Hoving 2014a; Wekker 2016; Boersma 2019; Vliet 2020), it becomes clear that the racializing Dutch governmental data systems and infrastructures are no “deviation” from an otherwise raceless Dutch culture. Furthermore, as shown by the childcare benefits scandal, and a welfare fraud detection system used in the municipality of Rotterdam<sup>81</sup> that recently came to light, the case studies discussed in this dissertation are no exceptions at all, but rather exemplary of racial attitudes and values currently embedded in many Dutch governmental data systems and infrastructures.

## **Governing Race-Ethnic Categorizations in a Datafied Society**

The results of this research clearly suggest that knowledge production about the relation between various societal phenomena and migration background, ethnicity, nationality, and race, should be organized in a better way. Furthermore, the results encourage us to think not only about categorization and clustering practices but also about the ways in which categorizations and clusters are impacted by their datafication. In the following section, I will point at directions in which we can think further about these topics. I will make a few suggestions for how choices regarding categorization, clustering, and datafied dissemination of race-ethnic categories can be made in a responsible manner. In this way, we are able to formulate how governmental institutions can create policies and practices for anti-racist value-driven datafied governance.

### **Classification and Race-ethnic Determination/Self-Determination**

The case studies of this dissertation raise questions about the non-consensual manner in which people are assigned a migration background by the Dutch government and the normalization of the dissemination of such information via municipal registrations and CBS APIs. As a result of this combination, data points concerning migration background are available for processes of technologically mediated racialization regardless of the intentions for reuse. One of the ways in which people could be given more agency in this process is by making migration background or race-ethnicity a matter of choice. Similar to the increasing possibilities for people to change their assigned sex it should not be too difficult to create possibilities for people to communicate to which race-ethnic group or groups they feel they belong. This could be done either through the

---

<sup>81</sup> There is no academic work on this system yet, but the list of indicators used in this algorithm can be found in the attachments of a letter from Richard Moti, an alderman of the city of Rotterdam, that was made publicly available after a request from journalists (see Tromp, Delhaas, and Davidson 2021)

existing infrastructure that includes the municipal administration and CBS, or via a process of census taking as is the case in the United States.

One of the benefits of the combination of race-ethnic self-determination and the possibility to choose multiple race-ethnic groups is that race-ethnic categories would be less essentialist and more suitable for monitoring processes of inclusion and exclusion. Opting for a combination of multiple migration backgrounds would allow for more detailed information about people and a more nuanced, less essentialist view on what is meant by particular categories. Furthermore, the main goal of the use of historically stigmatized categories should ultimately be their own abolishment; categories should work to make themselves obsolete. Allowing people to communicate those personal characteristics on the basis of which they feel discriminated against, provides more detailed and in-depth information about exclusionary processes. As such, self-assigned categories are, at least in theory, more suitable for anti-racist and anti-discriminatory purposes.

### **Clustering, Labeling, and Situatedness**

The use of state sanctioned, centralized, and datafied clusters and labels implies a distinct assumption about knowledge production and its perceived objectivity that does not take into account the social, cultural, and historical aspects that are entangled with the act of knowledge and world creation. As advised by WRR in 2016, we should, therefore, aim not for generalizable centralized categories, clusters and labels, but rather for categorization systems tailored to the problems they are trying to solve. For example, for an initiative monitoring racial discrimination on the basis of skin colour in hiring practices, a dataset based on migration background, or clusters on the basis of language, are an ill fit. Ideally, categorizations and clusters should, therefore, be as similar to the phenomenon being investigated as possible.

Colonial or racializing connotations should be avoided in assigning labels to categories used in government statistics and policy. In the Netherlands, words like “blank” or “Western” have been largely abolished in the past decade for reasons that in the instrumental approach would fall to technologically mediated racialization. There is, however, more room for improvement. From an epistemological perspective, labels of categories should, ideally, avoid presenting as self-evident, natural, or essentialist. This can be done by always drawing attention to the constructedness of any category by placing it between quotation marks and avoiding phrases such as “Daniel *is* a person with a migration background”. Instead, the act of categorization could be made explicit by stating that *in the context of a particular research question or policy initiative* people with certain characteristics *are categorized as* “people with a migration background.” Such a use of categorizations makes clear that research questions, situations, and categorizations are entangled entities that should not crossover to other domains without careful consideration.



Through an active and continuous focus on contextualization, it becomes possible to anticipate and counter the decontextualization inherent in datafication.

To aid in practices like these, in 2021, the Dutch Council for Government Policy (WRR) suggested four criteria to test the suitability of categories and clusters: 1) legitimacy; they should aid in a legitimate goal, 2) functionality; they should aid in achieving this goal, 3) subsidiarity; the goal can not be achieved by other means, 4) proportionality; benefits of categorization and clustering should outweigh drawbacks (see R. Jennissen, Bovenkerk, and Engbersen 2021, 14). While the advice emphasizes that these considerations should be made for specific situations and research questions, it does imply that the abovementioned criteria can be considered objectively. The problem with this assumption is that it does not take into account that Dutch society is not a level playing field, but rather an environment with various inequalities built in. As a result, perspective matters when assessing a new system for its legitimacy, functionality, subsidiarity, and proportionality. Especially when envisioned “benefits” and “drawbacks” are not carried equally by people from different race-ethnic groups. To account for such differences, we need a way to consider situatedness in categorization systems and their applications; perspective matters in knowledge production, and even more when such knowledge is disseminated in a datafied manner.

### **Datafication and Accountability - Reversing the Order**

As I have shown in the previous chapters, the path governmental data technologies take follows the order in which I presented the different analytical approaches to technologically mediated racialization. First comes instrumentalism, in which a government agency decides to implement a data technology with the idea to make its conduct more “data-driven” and therefore, as is the assumption, more “efficient”. As shown in the cases of the Leefbaarometer and the Crime Anticipation System, these ideas are largely driven by the rather technological solutionist “Big Data” discourse. In the next, epistemological, step, data is selected and combined on the basis of availability and sociocultural ideas about relevance. As shown, relevance and availability often overlap; when race-ethnic data is available, it is assumed to be relevant, and when race-ethnic data is considered relevant, it is usually available. In a final, more ontological step, racializing data systems are implemented, despite negative results or absence of proof for efficiency.<sup>82</sup>

---

<sup>82</sup> It should be noted that this does not only happen when data about race-ethnicity is involved. Other well-known examples of questionable governmental data-driven applications are Covid-19 contact tracing apps, of which the development was largely driven by technological solutionist discourse, rather than proof of function or efficiency (see L. Taylor 2021).

What is striking about the many racializing systems that have been implemented in practices of governance, fraud detection, and policing, is that they seem not to be consciously driven by ethical considerations or particular values, but rather by their sheer technological possibility. By taking the phrase “data-driven” rather literally, the produced values produce and reproduce race-ethnic inequalities. If we, as a society, want to turn this around, we have to quite literally reverse the order and work from ontology, via epistemology, towards an instrument. Only in this way can we replace technological solutionism, with anti-racist, ethnical, and value-driven governance.

### **Step 1: Ontology – What kind of world do we aim for?**

The ontological approach to race and data can be understood as a process of worldmaking. Working towards a future society where skin colour, migration background, and nationality are no longer influential factors for the life chances of groups and individuals, requires a vision of what such a world could look like. Article 1 of the Dutch constitution could be a suitable guiding principle in this process:

All persons in the Netherlands shall be treated equally in equal circumstances. Discrimination on the grounds of religion, belief, political opinion, race or sex or on any other grounds whatsoever shall not be permitted. (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties 2002, 5)

When we consider this article in Dutch law together with the main argument of this dissertation, it should be evident that race-ethnic data should never be used for policing, fraud detection, and surveillance practices, since such data can only reproduce existing racializing processes.

However, as addressed earlier (see Chapter 1), race-ethnic data can also be helpful in working towards a more equal future. Take for example the *Barometer Culturele Diversiteit*, or Barometer Cultural Diversity (BCD), an application created by CBS that measures the diversity in migration background of the personnel of large organizations (see van Schie forthcoming). This system was made on the basis of the very same infrastructure as the other case studies in this dissertation, including the racializing vocabulary of “western” and “non-Western”. Through the epistemological approach, we would conclude that the knowledge produced by BCD is haunted by the same histories of racism and colonialism discussed in this dissertation. Furthermore, the BCD conceptualization of diversity—solely in terms of migration backgrounds—is rather narrow. Through an ontological approach, however, we can see that there are reasons why the use of racialized categories in systems monitoring “diversity” might be warranted. The clustered migration backgrounds signified by the orientalist and colonial designations of “Western” and “non-Western” distinguish quite precisely the ideological demarcations between non-stigmatized migration backgrounds and stigmatized migration backgrounds in the context of the Netherlands. While it might seem counterintuitive, these clusters of categories, ideally with new

labels, are rather fitting for the task for which the BCD was created: measuring how people from historically stigmatized groups are doing in large organizations. The practice of using stigmatized categories and vocabulary in an attempt to subvert their effects is referred to in feminist and postcolonial discourses as the already mentioned “strategic essentialism” (see Eide 2016). In the words of postcolonial scholar Gayatri Spivak: “You pick up the universal that will give you the power to fight against the other side, and what you are throwing away by doing that is your theoretical purity” (Spivak in Spivak and Harasym 2014, 12). Such a pragmatic use of essentialism should, however, be treated with care, and only temporarily, to prevent the reification and naturalization of categories one eventually wants to get rid of.

While anti-racist data systems, such as the abovementioned BCD, do produce race-ethnicity in a way that follows the colonial and racist logic inherited by the categories they use, they do so in a way that aims to destroy the effects of these categories in organizations in the future. Nevertheless, it does remain important to realize that the results of BCD imply that the representation of people with different migration backgrounds should eventually reflect numbers that correspond with the values of institutions that truly value inclusion and diversity; it implies both change in organizational culture and the implementation of policies and initiatives to enact that change. To achieve such a future, we need clear values, a vision, a horizon to work towards.

## **Step 2: Epistemology—What Do We Need to Know?**

After appropriate values and goals are set by an organization, the next step to consider is what knowledge and perspectives are needed to achieve the aims that are being pursued. As a rule-of-thumb, it is advisable to refrain from using race-ethnic categorization in most investigations or policy situations where race-ethnicity is not already part of the research question. However, when an organization wants to know whether the representation of race-ethnic minority groups in their personnel reflects the percentages of these groups in the local population, it does make sense to try to gather data about the specific differences that seem to hinder people. Ideally, such investigations incorporate differences along multiple axes, since the intersection of different identity characteristics, such as race, gender, sexuality can result in specific forms of discrimination (see Crenshaw 1990).

Definitions of categories should be clearly defined and always be chosen in relation to the research questions. It is advised to continue to make explicit that categories are in no way natural or self-evident, for example by making the word “category”, or an equivalent, part of all descriptions. In some cases, it might even be possible to simply number categories or clusters rather than make their labels adhere to discriminatory or stigmatizing mainstream terms. Furthermore, normative categorizations can better be avoided. For example, the categorization system used by CBS distinguishes between people with a “Dutch background” and with a

“migration background”. While it might seem like a minor detail, this phrasing already suggests that in the case of a Dutch background, migration does not play a role. If we want to limit normative assumptions, it would make sense to begin to consider the Netherlands as just one migration background like any other, and only use this category in situations where its use is warranted.

Finally, it should be noted that most questions about oppression can usually be answered best by the people that are directly affected by it. When representation of a particular group in an organization is lacking this is usually not accidental, but rather the result of particular policies or a particular organizational culture. Finding out about the numbers of people through a datafied application does not explain anything about the reasons for inequality. In our research questions we should aim to not only find out what the current situation is, but also what the underlying forces are that create this situation. For example, in the case of fighting crime, creating a predictive policing system on the basis of personal characteristics does nothing to understand why people with those characteristics are at a higher risk to become a victim or perpetrator. In such instances, categorizations are not working towards their own obsolescence, but are, rather, exploiting societal inequalities. While these practices are often justified on the basis of an assumed “efficiency”, they are in stark contrast with the abovementioned first article of the Dutch constitution. On the basis of my findings in this dissertation, it has, therefore, become my strong conviction that such an exploitation of societal inequalities should have no place in contemporary datafied governance.

### **Step 3: Instrumentality – Picking the Right Tools for the Job**

When it is clear what kinds of knowledge are necessary in light of the envisioned goals, it is time to choose the appropriate technologies to produce this knowledge. In some cases, a data system or interactive map might be the best possible approach. In other situations, however, questionnaires or interviews might do. Take for example the previously mentioned case of the Dutch police. It has been well-reported since the 1980s that the police organization has problems with the recruitment and retention of police officers with a migration background (see de Ruijter 1998; Hart-Kemper and Nas 1998; Broekhuizen, Raven, and Driessen 2007). The reasons for these processes, which are considered to be a combination of work place discrimination, microaggressions and institutional racism, have been well investigated, and more importantly, without reference to any datafied methods (see Broekhuizen, Raven, and Driessen 2007; Mutsaers 2014; Çankaya 2017; Gowricharn and Çankaya 2017). When BCD was used to investigate the race-ethnic make-up of the Dutch national police in 2017, it again became clear that people with a “non-Western” migration background were well represented at the different police academies, but that they tended to quit working for the police in their first few years of service at a disproportionate rate compared to people with a Dutch migration background (see CBS 2017a). As a result of this, the police started an investigation into the reasons for this

discrepancy (see Politie Nederland 2020), seemingly disregarding the scholarly work that already exists about this topic. In this case, BCD did not add anything to the existing knowledge about the race-ethnic make up of the police, and more importantly, mere numbers cannot provide any solutions as I already explained. The point I am trying to make with this example is that in many policy situations that relate to race-ethnicity, ethnographic and other qualitative work can be much more informative than any datafied methods can be. Organizations should, therefore, always seriously consider whether the creation of a data system fits the knowledge and vision they aim for.

A final, very important step in cases where a datafied method is chosen as a fitting solution for a problem, is testing whether it works. While this might seem rather self-evident, the cases of the Rotterdam Act/Leefbaarometer combination, and CAS show that proof for claims of efficiency is skipped rather easily when policymakers have a strong belief in future accomplishments. At this point it should be mentioned that the data and algorithm used in the childcare benefits scandal too, was not only problematic because of its bias towards people with a migration background and/or a dark skin colour, but also because the process did not work. The victims in this case were not only accused of fraud, but importantly, *wrongfully* accused of fraud. Efficiency and accuracy of algorithmic systems should, therefore, never simply be assumed, but always tested for requirements that are set in advance. In the case of failure, systems should be redesigned and tested again. In this way systems can only be rolled out when necessary requirements are met, or, when they are not met, when sufficient reasons are given for its assumed trustworthiness.

The reason for finishing a process by choosing and testing instruments is that, when the proper order is followed, the systems and solutions being tested are anti-racist systems. In this way it becomes much more difficult to accidentally implement working solutions for policy problems that are featuring one or more forms of technologically mediated racialization. Only by systematically working through the different levels will we be able to have fruitful political discussions about how to create an equal and just datafied society in the Netherlands and provide the conditions for people of all migration backgrounds, nationalities, and other axes of difference to take part in it.

### **Further Research on Technologically Mediated Racialization**

Although I developed my conceptual framework of technologically mediated racialization for the analysis of race-ethnicity in data systems and infrastructures in the Netherlands, I strongly feel that it is suitable for the investigation of different technologies, infrastructures, and contexts as well. In the following paragraphs, I will outline some promising directions in which this research can be taken, and for which the groundwork is already being laid in critical data studies, postcolonial studies, and related disciplines.

First, in this dissertation, all case studies were connected to the same race-ethnic data ontology based on migration background. Another direction that could be explored with the concept of technologically mediated racialization is into data systems that take other human characteristics as their main focal point. For example, in a case like the DNA-portraits discussed in the introduction, race-ethnicity is constructed on the basis of genomics. As I briefly showed, this construction of Dutchness is not only connected to age-old discourses on blood, soil, and nation. DNA as a human trait available to be racialized only exists in relation to the technologies that make it visible (see also M'charek and van Oorschot 2020). In this sense, the existence of DNA-technology has an ontological imperative on a specific situated possibility for the construction of race. This technicity of race creates different possibilities for racialization than the race-ethnic categorization system investigated in the case studies of this dissertation. In a similar fashion, many other human characteristics are routinely used to construct and reify historically racialized differences between people. Technologies like fingerprinting, facial recognition, and gait recognition all measure physical characteristics of people and are mediating processes of racialization (see, for example, Pugliese 2007; Browne 2015; Vukov 2016; Leurs and Shepherd 2017; Madianou 2019). In each of these cases, technologically mediated racialization takes on a situated form that depends on the affordances of the technologies used, their purposes, and their geographic location. Investigating such cases through the three approaches to technologically mediated racialization may aid in a better understanding not only of the effects of racializing technologies, but also of the mediating process for which those technologies are selected.

A second possible direction of research is the comparison of similar technologies in different geographical and national contexts to highlight the relevance of their context and situatedness. Where the ontological approach might focus on the mediated racializing process in a somewhat technologically determinist way, the comparison of similar technologies in different contexts might highlight the socially constructed nature of the ways in which the technologies acquire their meaning and purpose. The combination of technologies in a different national and geographical context might also allow for comparisons that show not only the sociotechnical differences between contexts, but also their similarities. In contrast to the widely shared belief in the Netherlands that racism is mostly “an American problem” (Özdil 2014), the colonial histories *and* conceptions of race of countries like the United States, the Netherlands, France, and South Africa show many connections and overlaps (Stoler 1995; 2009; 2016). This raises the question if and how these histories and racial formations translate and interact differently with data technologies in these local contexts.

A third possible direction of research is into how the instrumental, epistemological, and ontological approaches to technologically mediated racialization can aid in better understanding the ways in which race-ethnic categorizations can be used for anti-racist and decolonial purposes.

Since the case studies in this dissertation were selected for their racializing use of race-ethnic categories, they tended to feature racialization at all levels of analysis. However, preliminary research (see van Schie, forthcoming) seems to indicate that in systems for the purpose of monitoring diversity or affirmative action, results might differ depending on the approach. It could be argued, that for purposes of anti-discrimination and affirmative action, strategic essentialism is a viable option. The ontological dimension of this strategy would not lie in the supposed reality of the essentialist categories, but, rather, in the future one is attempting to create. Ideally, eventually, Dutch organizations will become the meritocracies they are already pretending to be. Striving for such a goal can hardly be called racist, even if a system uses categories that would be considered forms of racialization in situations like policing of fraud detection. I would, therefore, argue that whether the use of race-ethnic categories should be considered a form of racialization depends strongly on the future it is helping to realize. While BCD does produce race-ethnicity in a way that follows the colonial and racist logic inherited by the categories it uses, it does so in a way that aims to paint a picture of the current state of affairs so that affirmative action can be taken based on the results. The categories used in BCD are there to make themselves obsolete, which is an entirely different purpose than that of categories used in systems such as the Leefbaarometer and the Crime Anticipation System from Chapters 4 and 5.

Finally, academics should invest in finding out how to implement research findings in policy practice by doing experiments and collaborations with government institutions and other societal partners. Inspiring work is already being done by institutes such as Data & Society in New York as well as the Data Justice Lab at the University of Cardiff. Furthermore, much closer to home, the Utrecht Data School produced the Data Ethics Decision Aid (DEDA), a tool that civil servants can use in the planning and design phase of any project that involves data (see Franzke, Muis, and Schäfer 2021), and the Impact Assessment Human Rights and Algorithms (IAMA), another tool that specifically checks whether or not an algorithmic system aligns with basic human rights. While the implementation of tools such as DEDA and IAMA is currently largely dependent on the willingness of individual civil servants, policymakers, and institutions, we should work towards a situation in which the use of tools like these are standard practice. Only by avoiding arbitrariness, can we, as a society, ensure that human rights are not an afterthought but a guiding principle in the operationalization of race-ethnic categories and algorithms. In this way, I hope that governments in general and Dutch governmental institutions in particular are able to work towards a more accountable and just datafied society that values all people equally.

## Bibliography

- Abraham, Itty. 2006. "The Contradictory Spaces of Postcolonial Techno-Science." *Economic and Political Weekly* 41 (3): 210–17.
- Adams, Vincanne. 2002. "Randomized Controlled Crime: Postcolonial Sciences in Alternative Medicine Research." *Social Studies of Science* 32 (5–6): 659–90. <https://doi.org/10.1177/030631270203200503>.
- Al Jaramani, Araa, Sandra Ponzanesi, and Gerwin van Schie. 2022. "'Girls Are like Glass': Situated Knowledges of Syrian Refugee Women on Datafication and Transparency." In *Interrogating Datafication: Towards a Praxeology of Data*, edited by Marcus Burkhardt, Daniela van Geenen, Carolin Gerlitz, Sam Hind, Timo Kaerlein, Danny Lämmerhirt, and Axel Volmar, 117–37. Media in Action. Bielefeld: transcript Verlag.
- Amnesty International. 2013. "Proactief Politieoptreden Vormt Risico Voor Mensenrechten. Etnisch Profileren Onderkennen En Aanpakken." Amsterdam: Amnesty International, Afdeling Nederland.
- Amoore, Louise. 2011. "Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times." *Theory, Culture & Society* 28 (6): 24–43. <https://doi.org/10.1177/0263276411417430>.
- Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Revised ed. London: Verso.
- Anderson, Chris. 2008. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired*. 2008. <https://www.wired.com/2008/06/pb-theory/>.
- Anderson, Warwick. 2002. "Introduction: Postcolonial Technoscience." *Social Studies of Science* 32 (5/6): 643–58.
- Andersson Elffers Felix. 2014. "Onderzoek naar de betekenis van integrale bevraging voor het operationele politiewerk." gv262. Utrecht: Politieacademie.
- Andrejevic, Mark, Lina Dencik, and Emiliano Treré. 2020. "From Pre-Emption to Slowness: Assessing the Contrasting Temporalities of Data-Driven Predictive Policing." *New Media & Society* 22 (9): 1528–44. <https://doi.org/10.1177/1461444820913565>.
- Appadurai, Arjun. 2006. *Fear of Small Numbers*. Durham NC, United States: Duke University Press. <https://doi.org/10.1215/9780822387541-004>.
- Appiah, Kwame Anthony. 1991. "Is the Post- in Postmodernism the Post- in Postcolonial?" *Critical Inquiry* 17 (2): 336–57.
- Arora, Payal. 2016. "Bottom of the Data Pyramid: Big Data and the Global South." *International Journal of Communication* 10 (2016): 1681–99.
- . 2019. *The Next Billion Users*. Harvard University Press.
- Ashcroft, Bill, Gareth Griffiths, and Helen Tiffin. 1989. *The Empire Writes Back: Theory and Practice in Post-Colonial Literatures*. 2nd ed. London & New York: Routledge.
- , eds. 1995. *The Post-Colonial Studies Reader*. London & New York: Routledge.
- Autoriteit Persoonsgegevens. 2020. "Belastingdienst/Toeslagen: De Verwerking van de Nationaliteit van Aanvragers van Kinderopvangtoeslag." Onderzoeksrapport z2018-2245. Den Haag: Autoriteit Persoonsgegevens.



- [https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek\\_belastingdienst\\_kinderopvangtoeslag.pdf](https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/onderzoek_belastingdienst_kinderopvangtoeslag.pdf).
- Banton, Michael. 1969. "Social Aspects of the Race Question." In *Four Statements on the Race Question*, edited by UNESCO, 9–16. Paris: Unesco. <http://unesdoc.unesco.org/images/0012/001229/122962eo.pdf>.
- Bauman, Zygmunt. 2007. *Modernity and the Holocaust*. Cambridge: Polity Press.
- Beck, Charlie. 2009. "Predictive Policing: What Can We Learn from Wal-Mart and Amazon about Fighting Crime in a Recession?" *The Police Chief*, November 2009. [https://controverses.minesparis.psl.eu/public/promo16/promo16\\_G16/acmst373ethics.weebly.com/uploads/2/9/6/2/29626713/police-chief-magazine.pdf](https://controverses.minesparis.psl.eu/public/promo16/promo16_G16/acmst373ethics.weebly.com/uploads/2/9/6/2/29626713/police-chief-magazine.pdf).
- Benjamin, Ruha. 2016. "Innovating Inequity: If Race Is a Technology, Postracialism Is the Genius Bar." *Ethnic and Racial Studies* 39 (13): 2227–34. <https://doi.org/10.1080/01419870.2016.1202423>.
- . 2019. *Race after Technology: Abolitionist Tools for the New Jim Code*. Cambridge: Polity.
- Berkhout, Frans. 2002. "Technological Regimes, Path Dependency and the Environment." *Global Environmental Change* 12 (1): 1–4. [https://doi.org/10.1016/S0959-3780\(01\)00025-5](https://doi.org/10.1016/S0959-3780(01)00025-5).
- Betancur, John J., and Cedric Herring, eds. 2013. *Reinventing Race, Reinventing Racism*. Studies in Critical Social Sciences. Leiden & Boston: Brill.
- Bezemer, W., and A. Leerkes. 2021. *Oververtegenwoordiging verder ontcijferd: Een kwantitatief onderzoek naar sociale verschillen in verdenkingskans en zelfgerapporteerd crimineel gedrag*. Politie & Wetenschap. Den Haag: Sdu Uitgevers.
- Bezuidenhout, Louise M., Sabina Leonelli, Ann H. Kelly, and Brian Rappert. 2017. "Beyond the Digital Divide: Towards a Situated Approach to Open Data." *Science and Public Policy* 44 (4): 464–75. <https://doi.org/10.1093/scipol/scw036>.
- Bhabha, Homi K. 2010. *The Location of Culture*. London: Routledge.
- Boersma, Sanne. 2019. "Dissociating Society: Knowledge, Affect and Performativity in Immigrant Integration Monitoring." Dissertation, Rotterdam: Erasmus University Rotterdam.
- Boersma, Sanne, and Willem Schinkel. 2015. "Imagining Society: Logics of Visualization in Images of Immigrant Integration." *Environment and Planning D: Society and Space* 33 (6): 1043–62. <https://doi.org/10.1177/0263775815598153>.
- Bonjour, Saskia, and Laura Block. 2016. "Ethnicizing Citizenship, Questioning Membership. Explaining the Decreasing Family Migration Rights of Citizens in Europe." *Citizenship Studies* 20 (6–7): 779–94. <https://doi.org/10.1080/13621025.2016.1191429>.
- Boogers, M.J.G.J.A., B. Becker, and P.W. Tops. 2005. "Experiment burgerplatform VROM: lessen en leerervaringen voor burgerparticipatie en beleidsinnovatie." Tilburg: University of Tilburg.
- Bovenkerk, Frank. 1978a. "Enkele Suggesties Voor Het Bestrijden van Discriminatie." In *Omdat Zij Anders Zijn: Patronen van Rasdiscriminatie in Nederland*, edited by Frank Bovenkerk, 189–202. Amsterdam: Boom Koninklijke Uitgevers.
- . 1978b. "Rasdiscriminatie in Nederland." In *Omdat Zij Anders Zijn: Patronen van Rasdiscriminatie in Nederland*, edited by Frank Bovenkerk, 9–30. Amsterdam: Boom Koninklijke Uitgevers.

- . 1984. “Rassen of Klassen? De Politieke Economie van de Gastarbeid.” *Intermediair* 20 (47): 23.
- . 2000. “The Other Side of the Anne Frank Story: The Dutch Role in the Persecution of the Jews in World War Two.” *Crime, Law & Social Change* 34 (3): 237–58.
- Bovens, Mark, Meike Bokhorst, Roel Jennissen, and Godfried Engbersen. 2016. “Migratie En Classificatie: Naar Een Meervoudig Migratie-Idioom.” Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid.
- Bovens, Mark, Roel Jennissen, Meike Bokhorst, and Godfried Engbersen. 2021. “Afscheid van westers en niet-westers: Naar meervoudige indelingen van herkomstgroepen.” WRR-policy brief 8. Den Haag: WRR.
- Bowker, Geoffrey C. 2013. “Data Flakes: An Afterword to ‘Raw Data’ Is an Oxymoron.” In *“Raw Data” Is an Oxymoron*, edited by Lisa Gitelman, 167–72. Infrastructures Series. Cambridge, Massachusetts ; London, England: The MIT Press.
- Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things out: Classification and Its Consequences*. Cambridge, MA: MIT press.
- boyd, danah, and Kate Crawford. 2012. “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.” *Information, Communication & Society* 15 (5): 662–79. <https://doi.org/10.1080/1369118X.2012.678878>.
- Braddock, Clarence H. 2021. “Racism and Bioethics: The Myth of Color Blindness.” *The American Journal of Bioethics* 21 (2): 28–32. <https://doi.org/10.1080/15265161.2020.1851812>.
- Braidotti, Rosi. 2013. *The Posthuman*. Cambridge, UK, Malden, MA: Polity Press.
- Brayne, Sarah. 2021. *Predict and Surveil: Data, Discretion, and the Future of Policing*. New York (N.Y.): Oxford University Press.
- Brienen, Rebecca P. 2014. “Types and Stereotypes: Zwarte Piet and His Early Modern Sources.” In *Dutch Racism*, edited by Philomena Essed and Isabel Hoving, 179–200. Amsterdam: Rodopi.
- Brion, Fabienne. 2004. “Een Nederlandse Eigenaardigheid. Onderzoek Naar Criminaliteit van Allochtonen.” In *Tijdschrift Voor Criminologie*, edited by P. Kleemans, P. van der Laan, and P. Ponsaers, Jubileumuitgave, 66–76. Amsterdam: SISWO.
- . 2005. “Criminaliteit en Immigratie: Stand van Zaken.” In *Culturele diversiteit en samenleven in Brussel en België*, edited by Marc Swyngedouw, Pascal Delwit, and Andrea Rea. Leuven: ACCO.
- Broekhuizen, J., J. Raven, and F.M.H.M. Driessen. 2007. *Positie En Expertise van de Allochtone Politiedewerker. Op Weg Naar Een Volwaardige Plaats Binnen de Politie*. Politiewetenschap 37. Utrecht: Bureau Driessen.
- Brown, Mark B. 2015. “Politicizing Science: Conceptions of Politics in Science and Technology Studies.” *Social Studies of Science* 45 (1): 3–30. <https://doi.org/10.1177/0306312714556694>.
- Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press.
- Buikema, Rosemarie L. 2009. “The Arena of Imaginings: Sara Bartmann and the Ethics of Representation.” In *Doing Gender in Media, Art and Culture*, edited by Rosemarie L. Buikema and Iris Van der Tuin. London: Routledge.

- . 2012. "Monumental Dresses: Coming to Terms with Racial Repression." In *Teaching Race with a Gendered Edge*, edited by Brigitte Hipfl and Kristín Loftsdóttir, 43–60. Budapest & New York: Central European University Press.
- Buolamwini, Joy, and Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of Machine Learning Research*, 77–91.
- Çankaya, Sinan. 2012. *De Controle van Marsmannetjes En Ander Schorriemorrie. Het Beslissingsproces Tijdens Proactief Politiewerk*. Amsterdam: Boom Lemma.
- . 2015. "De Politiële Surveillance van Ras En Etniciteit." In *Ethnic Profiling En Interne Diversiteit Bij de Politie*, edited by L.G. Moor, J. Janssen, M. Easton, and A. Verhage, 2:13–33. Cahiers Politiestudies 35. Antwerpen: Maklu-Uitgevers.
- . 2017. "The Racialization of Ethnic Minority Police Officers and Researchers: On Positionality and (Auto)Ethnographic Fieldwork." *European Journal of Policing Studies* 5 (Special Issue [1]): 119–32.
- . 2020. "Geopolicing Race, Gender, and Class: How the Police Immobilise Urban Allochthones." *Antipode* 52 (3): 702–21. <https://doi.org/10.1111/anti.12613>.
- Carvalho, Anabela. 2008. "Media(Ted) Discourse and Society: Rethinking the Framework of Critical Discourse Analysis." *Journalism Studies* 9 (2): 161–77. <https://doi.org/10.1080/14616700701848162>.
- CBS. 1901. *Uitkomsten Der Achtste Tienjaarlijksche Volkstelling in Het Koninkrijk Der Nederlanden Gebouden Op de Een En Dertigsten December 1899. Bijdragen Tot de Statistiek van Nederland 3*. 's Gravenhage: Centraal Bureau voor de Statistiek.
- . 1911. *Uitkomsten Der Negende Tienjaarlijksche Volkstelling in Het Koninkrijk Der Nederlanden Gebouden Op de Een En Dertigsten December 1909. Bijdragen Tot de Statistiek van Nederland 154*. 's Gravenhage: Centraal Bureau voor de Statistiek.
- . 1922. *Volkstelling 31 December 1920*. Statistiek van Nederland 352. 's Gravenhage: Centraal Bureau voor de Statistiek.
- . 1933. *Volkstelling 31 December 1930. Vol. 2*. Statistiek van Nederland. 's Gravenhage: Algemeene Landsdrukkerij.
- . 1954. *12e Volkstelling Annex Woningtelling 31 Mei 1947*. Rijks- En Provinciale Cijfers 1a. Utrecht: Centraal Bureau voor de Statistiek.
- . 1966. *13e Algemene Volkstelling 31 Mei 1960*. Hilversum: Centraal Bureau voor de Statistiek.
- . 1978. *14e Algemene Volkstelling Annex Woningtelling 28 Februari 1971. Vol. 1b: niet-nederlandse nationaliteiten*. 's Gravenhage: Staatsuitgeverij.
- . 2014. *Dutch Census 2011 - Analysis and Methodology*. The Hague/Heerlen: Statistics Netherlands.
- . 2016. "Wie zijn de derde generatie?" Webpagina. 2016. <http://www.cbs.nl/nl-nl/nieuws/2016/47/wie-zijn-de-derde-generatie->
- . 2017a. "Migratieachtergrond Werknemers Politie Nederland 2017: Pilot Barometer Culturele Diversiteit." CBS. [https://www.cbs.nl/-/media/\\_excel/2019/05/190117\\_eigen\\_data\\_politie\\_nederland.xlsx](https://www.cbs.nl/-/media/_excel/2019/05/190117_eigen_data_politie_nederland.xlsx).

- . 2017b. “Totale In- En Uitvoer Naar Land.” CBS StatLine. July 8, 2017. <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=81271NED&D1=1&D2=a&D3=l&HDR=T%2CG2&STB=G1&VW=T>.
- . 2017c. “Bevolking; Generatie, Geslacht, Leeftijd En Herkomstgroepering, 1 Januari.” CBS StatLine. July 18, 2017. <http://statline.cbs.nl/statweb/publication/?dm=slnl&pa=37325&d1=0&d2=a&d3=0&d4=0&d5=a&d6=l&hdr=g2>.
- . 2018a. “Begrippen.” Webpagina. 2018. <http://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen>.
- . 2018b. “Explosieve groei gebruik open data CBS.” Webpagina. Centraal Bureau voor de Statistiek. 2018. <https://www.cbs.nl/nl-nl/corporate/2018/31/explosieve-groei-gebruik-open-data-cbs>.
- . 2018c. “Volkstelling 1971 wordt media-event.” Webpagina. 2018. <http://www.cbs.nl/nl-nl/onze-diensten/publicaties/het-licht-van-de-statistiek/1960-1980/volkstelling-1971-wordt-media-event>.
- . 2019. “Wat verstaat het CBS onder een allochtoon?” Webpagina. Centraal Bureau voor de Statistiek. 2019. <https://www.cbs.nl/nl-nl/faq/specifiek/wat-verstaat-het-cbs-onder-een-allochtoon->.
- . 2021a. “Bevolkingsstatistiek.” Webpagina. Centraal Bureau voor de Statistiek. 2021. <https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/bevolkingsstatistiek>.
- . 2021b. “Hoeveel allochtonen wonen er in Nederland?” Webpagina. Centraal Bureau voor de Statistiek. 2021. <https://www.cbs.nl/nl-nl/faq/specifiek/hoeveel-allochtonen-wonen-er-in-nederland->.
- . 2021c. “StatLine - Bevolking; geslacht, leeftijd, generatie en migratieachtergrond, 1 januari.” 2021. <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37325/table?fromstatweb>.
- . 2021d. “WoonOnderzoek Nederland (WoON).” Webpagina. Centraal Bureau voor de Statistiek. 2021. <https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/woononderzoek-nederland--woon-->.
- . 2021e. “Uithuisplaatsingen onder gedupeerden toeslagenaffaire.” Webpagina. Centraal Bureau voor de Statistiek. October 18, 2021. <https://www.cbs.nl/nl-nl/maatwerk/2021/42/uithuisplaatsingen-onder-gedupeerden-toeslagenaffaire>.
- Chemero, Anthony. 2003. “An Outline of a Theory of Affordances.” *Ecological Psychology* 15 (2): 181–95.
- Cheney-Lippold, John. 2016. “Jus Algoritmi: How the National Security Agency Remade Citizenship.” *International Journal of Communication* 10 (0): 22.
- . 2017. *We Are Data: Algorithms and the Making of Our Digital Selves*. New York: New York University Press.
- Chernilo, Daniel. 2006. “Social Theory’s Methodological Nationalism: Myth and Reality.” *European Journal of Social Theory* 9 (1): 5–22.

- Chun, Wendy Hui Kyong. 2009. "Introduction: Race and/as Technology; or, How to Do Things to Race." *Camera Obscura: Feminism, Culture, and Media Studies* 24 (1 (70)): 7–35. <https://doi.org/10.1215/02705346-2008-013>.
- . 2021. *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. Cambridge, Massachusetts London, England: The MIT Press.
- Coleman, Beth. 2009. "Race as Technology." *Camera Obscura: Feminism, Culture, and Media Studies* 24 (1 (70)): 177–207. <https://doi.org/10.1215/02705346-2008-018>.
- Connor, Walker. 1994. *Ethnonationalism*. Princeton, NJ: Princeton University Press.
- Controle Alt Delete. 2019. "Monitor etnisch profileren." Controle Alt Delete. 2019. <https://controlealtdelete.nl/monitor>.
- Cornelis, B. 1990. "Migratie Naar Nederland." In *Etnische Minderheden in Nederland*, edited by H. B. Entzinger and P. J. J. Stijnen, 13–30. Amsterdam & Heerlen: Boom Meppel & Open Universiteit.
- Couldry, Nick. 2017. "The Myth of Big Data." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 235–40. Amsterdam: Amsterdam University Press.
- Couldry, Nick, and Ulises Mejias. 2019a. *The Cost of Connection*. Stanford, CA: Stanford University Press.
- Couldry, Nick, and Ulises A. Mejias. 2019b. "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject." *Television & New Media* 20 (4): 336–49. <https://doi.org/10.1177/1527476418796632>.
- Crenshaw, Kimberle. 1990. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43 (6): 1241–1300.
- Creusen, Joep. 2014. "Open Data NL is hinken op twee benen." *Computable*. August 28, 2014. <https://www.computable.nl/artikel/opinie/datamanagement/5155816/1509029/open-data-nl-is-hinken-op-twee-benen.html>.
- Croes, Marnix. 2006. "The Holocaust in the Netherlands and the Rate of Jewish Survival." *Holocaust and Genocide Studies* 20 (3): 474–99.
- Dalton, Craig M., Linnet Taylor, and Jim Thatcher. 2016. "Critical Data Studies: A Dialog on Data and Space." *Big Data & Society* 3 (1): 1–9. <https://doi.org/10.1177/2053951716648346>.
- Dalton, Craig M., and Jim Thatcher. 2014. "What Does a Critical Data Studies Look like and Why Do We Care?" *Society & Space* (blog). May 12, 2014. <http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/>.
- Daniels, Jessie. 2013. "Race and Racism in Internet Studies: A Review and Critique." *New Media & Society* 15 (5): 695–719. <https://doi.org/10.1177/1461444812462849>.
- Darré, R. Walther. 1930. *Neuadel Aus Blut Und Boden*. München: J. F. Lehmanns Verlag.
- David, Paul A. 1985. "Clio and the Economics of QWERTY." *The American Economic Review* 75 (2): 332–37.
- Davis, Jenny L, and Nathan Jurgenson. 2014. "Context Collapse: Theorizing Context Collusions and Collisions." *Information, Communication & Society* 17 (4): 476–85. <https://doi.org/10.1080/1369118X.2014.888458>.

- De Villiers, John Abraham Jacob. 1896. *The Transvaal*. London: Chatto & Windus. <http://archive.org/details/transvaal00devi>.
- Dekker, Paul, and Peter Ester. 1996. "Depillarization, Deconfessionalization, and de-Ideologization: Empirical Trends in Dutch Society 1958-1992." *Review of Religious Research* 37 (4): 325–41.
- Denton, Nancy A. 1997. "Racial Identity and Census Categories: Can Incorrect Categories Yield Correct Information." *Law & Inequality* 15: 83–97.
- Derrida, Jacques. 1981. *Positions*. Translated by Alan Bass. Chicago: University of Chicago Press.
- . 1982. *Margins of Philosophy*. Translated by Alan Bass. New York, NY: The Harvester Press.
- Desrosières, Alain. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- D'Ignazio, Catherine, and Lauren F. Klein. 2020. *Data Feminism*. <Strong> Ideas Series. Cambridge, MA and London: The MIT Press.
- Dijk, Teun van. 2015. "Critical Discourse Analysis." In *The Handbook of Discourse Analysis*, edited by Deborah Tannen, Heidi E. Hamilton, and Deborah Schiffrin, Second Edition, 1:466–85. Malden, MA: Wiley-Blackwell.
- Dijstelbloem, Huub. 2021. *Borders as Infrastructure: The Technopolitics of Border Control*. Infrastructures. Cambridge, MA: MIT Press.
- Dijstelbloem, Huub, and Albert Meijer, eds. 2011. *Migration and the New Technological Borders of Europe*. Migration, Minorities and Citizenship. Houndmills, Basingstoke & Hampshire: Palgrave Macmillan.
- Dixon-Román, Ezekiel. 2016. "Algo-Ritmo: More-Than-Human Performative Acts and the Racializing Assemblages of Algorithmic Architectures." *Cultural Studies ↔ Critical Methodologies* 16 (5): 482–90. <https://doi.org/10.1177/1532708616655769>.
- . 2017. "Toward a Hauntology on Data: On the Sociopolitical Forces of Data Assemblages." *Research in Education* 98 (1): 44–58. <https://doi.org/10.1177/0034523717723387>.
- Doeleman, Reinder, René Melchers, Lauren Waardenburg, and Dick Willems. 2019. "3 misverstanden over predictive policing." *Websitevoordepolitie* (blog). November 21, 2019. <https://www.websitevoordepolitie.nl/3-misverstanden-over-predictive-policing/>.
- Drenth, Anthonie, and Ronald van Steden. 2017. "Ervaringen van straatagenten met het Criminaliteits Anticipatie Systeem." *Tijdschrift voor de Politie* 79 (3): 6–10.
- Driessen, F.M.H.M., F. Duursma, and J. Broekhuizen. 2014. "De Ontwikkeling van de Criminaliteit van Rotterdamse Autochtone En Allochtone Jongeren van 12 to 18 Jaar: De Rol van Achterstanden, Ouders, Normen En Vrienden." Apeldoorn and Utrecht: Politie & Wetenschap and Bureau Driessen.
- Drucker, Johanna. 2011. "Humanities Approaches to Graphical Display." *Digital Humanities Quarterly* 5 (1): 1–21.
- Dulk, C.J. den, and J.G.S.J. van Maarseveen. 1999. "Volkstellingen 1795-1971. De Ontwikkeling van Beleid En Methode van Onderzoek." In *Een Eeuw Statistieken: Historisch-Methodologische Schetsen van de Nederlandse Officiële Statistieken in de Twintigste Eeuw*, edited by B. Erwich and J.G.S.J. van Maarseveen, 329–66. Amsterdam: CBS.
- Eede, Yoni van den. 2011. "In Between Us: On the Transparency and Opacity of Technological Mediation." *Foundations of Science* 16 (2–3): 139–59. <https://doi.org/10.1007/s10699-010-9190-y>.

- Eide, Elisabeth. 2016. "Strategic Essentialism." In *The Wiley Blackwell Encyclopedia of Gender and Sexuality Studies*, 1–2. Hoboken, New Jersey: John Wiley & Sons. <https://doi.org/10.1002/9781118663219.wbegss554>.
- Eijkman, Quirine. 2010. "Has the Genie Been Let Out of the Bottle - Ethnic Profiling in the Netherlands." *Public Space: The Journal of Law and Social Justice* 5: 1–21.
- El-Tayeb, Fatima. 2011. *European Others: Queering Ethnicity in Postnational Europe*. Difference Incorporated. Minneapolis: University of Minnesota Press.
- Emigh, Rebecca Jean, Dylan Riley, and Patricia Ahmed. 2016a. *Antecedents of Censuses from Medieval to Nation States*. New York, NY: Palgrave Macmillan.
- . 2016b. *Changes in Censuses from Imperialist to Welfare States*. New York, NY: Palgrave Macmillan. <https://doi.org/10.1057/9781137485069>.
- Emirbayer, Mustafa, and Matthew Desmond. 2012. "Race and Reflexivity." *Ethnic and Racial Studies* 35 (4): 574–99. <https://doi.org/10.1080/01419870.2011.606910>.
- Enderson, Mary C., and David D. Selover. 2021. "The US Failure to Adopt the Metric System: The High Cost of Teaching the English System." *International Journal of Mathematical Education in Science and Technology* 0 (0): 1–16. <https://doi.org/10.1080/0020739X.2021.1890247>.
- Ensel, Remco. 2013. "Dutch Face-ism. Portrait Photography and Völkisch Nationalism in the Netherlands." *Fascism* 2 (1): 18–40. <https://doi.org/10.1163/22116257-00201009>.
- Es, Karin van, and Mirko Tobias Schäfer. 2017. "New Brave World." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 13–22. Amsterdam: Amsterdam University Press.
- Esmeijer, Loek, and Margreet Luning. 1978. "Surinamers in de Ogen van de Amsterdamse Politie." In *Omdat Zij Anders Zijn: Patronen van Rasdiscriminatie in Nederland*, edited by Frank Bovenkerk, 136–65. Amsterdam: Boom Koninklijke Uitgevers.
- Espeland, Wendy Nelson, and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology* 49 (03): 401. <https://doi.org/10.1017/S0003975609000150>.
- Essed, Philomena. 1991. *Understanding Everyday Racism: An Interdisciplinary Theory*. London and New York: Sage Publications.
- Essed, Philomena, and Isabel Hoving, eds. 2014a. *Dutch Racism*. Amsterdam: Rodopi.
- . 2014b. "Innocence, Smug Ignorance, Resentment: An Introduction to Dutch Racism." In *Dutch Racism*, edited by Philomena Essed and Isabel Hoving, 9–30. Amsterdam: Rodopi.
- Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. First Edition. New York, NY: St. Martin's Press.
- Fairclough, Norman. 2013. *Critical Discourse Analysis: The Critical Study of Language*. London: Routledge.
- Fajth, Veronika, and Özge Bilgili. 2020. "Beyond the Isolation Thesis: Exploring the Links between Residential Concentration and Immigrant Integration in the Netherlands." *Journal of Ethnic and Migration Studies* 46 (15): 3252–76. <https://doi.org/10.1080/1369183X.2018.1544067>.
- Fanon, Frantz. 2008. *Black Skin, White Masks*. Translated by Richard Philcox. New York, NY: Grove Press.

- Fassin, Didier. 2011. "Policing Borders, Producing Boundaries. The Governmentality of Immigration in Dark Times." *Annual Review of Anthropology* 40 (1): 213–26. <https://doi.org/10.1146/annurev-anthro-081309-145847>.
- . 2013. *Enforcing Order: An Ethnography of Urban Policing*. Cambridge & Malden, MA: Polity.
- Feenberg, Andrew. 1999. *Questioning Technology*. London & New York: Routledge.
- Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York, NY: Random House, Inc.
- . 1978a. *The Care of the Self*. Vol. 3. The History of Sexuality. New York: Pantheon Books.
- . 1978b. *The History of Sexuality: An Introduction*. Translated by Robert Hurley. Vol. 1. The History of Sexuality. New York: Pantheon Books.
- . 1980. *Power/Knowledge: Selected Interviews and Other Writings by Michel Foucault 1972–1977*. New York: Pantheon Books.
- . 1988. *The Use of Pleasure*. Translated by Robert Hurley. Vol. 2. The History of Sexuality. New York: Vintage Books.
- . 2006. *History of Madness*. Edited by Jean Khalifa. Translated by Jean Khalifa and Jonathan Murphy. London & New York: Routledge.
- . 2007. *Security, Territory, Population: Lectures at the Collège de France, 1977–78*. Translated by Graham Burchell. London and New York: Palgrave Macmillan.
- . 2013. "Right of Death and the Power of Life." In *Biopolitics: A Reader*, edited by Timothy C. Campbell and Adam Sitze, 41–60. Durham: Duke University Press.
- 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>.
- Frijtag Drabbe Künzel, Geraldien von. 2017. "'Germanje': Dutch Empire-Building in Nazi-Occupied Europe." *Journal of Genocide Research* 19 (2): 240–57. <https://doi.org/10.1080/14623528.2017.1313521>.
- Gambert, Iselin, and Tobias Linné. 2018. "From Rice Eaters to Soy Boys: Race, Gender, and Tropes of 'Plant Food Masculinity.'" *Animal Studies Journal* 7 (2): 129–79. <https://doi.org/10.2139/ssrn.3298467>.
- Gandy, Oscar H. 2016. *Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage*. London: Routledge.
- Gangadharan, Seeta Peña, and Jędrzej Niklas. 2019. "Decentering Technology in Discourse on Discrimination." *Information, Communication & Society* 22 (7): 882–99. <https://doi.org/10.1080/1369118X.2019.1593484>.
- Geenstijl. 2017. "Allochtoon-O-Meter." Allochtoon-O-Meter. 2017. [www.geenstijl.nl/paginas/mijnburen](http://www.geenstijl.nl/paginas/mijnburen).
- Gemeente Rotterdam. 2021. "Leefbaarheid & Veiligheid." Onderzoek 010. 2021. <https://onderzoek010.nl/dashboard/onderzoek010/leefbaarheid-en-veiligheid/>.
- Gemeente Utrecht. 2021. "Over de Monitor Leefbaarheid." Monitor Leefbaarheid. 2021. <https://utrecht.incijfers.nl/dashboard/monitor-leefbaarheid/over-de-monitor-leefbaarheid>.



- Gent, Wouter van, 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>.
- Ghorashi, Halleh. 2014. "Racism and 'the Ungrateful Other' in the Netherlands." In *Dutch Racism*, edited by Philomena Essed and Isabel Hoving, 101–16. Amsterdam: Rodopi.
- Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Gilroy, Paul. 1998. "Race Ends Here." *Ethnic and Racial Studies* 21 (5): 838–47. <https://doi.org/10.1080/014198798329676>.
- Gitelman, Lisa, ed. 2013. *"Raw Data" Is an Oxymoron*. Infrastructures Series. Cambridge, Massachusetts ; London, England: The MIT Press.
- Gordon, Avery F. 1997. *Ghostly Matters: Haunting and the Sociological Imagination*. Minneapolis & London: University of Minnesota Press.
- Gouweleeuw, José, and Carel Harmsen. 2004. "Foreigners at Work." In *The Dutch Virtual Census of 2001. Analysis and Methodology*, edited by Eric Schulte Nordholt, Marijke Hartgers, and Rita Gircour, 97–118. Voorburg/Heerlen: Statistics Netherlands.
- Gowricharn, Ruben, and Sinan Çankaya. 2017. "Policing the Nation: Acculturation and Street-Level Bureaucrats in Professional Life." *Sociology* 51 (5): 1101–17. <https://doi.org/10.1177/0038038515601781>.
- Grapperhaus, F.B.J, S. Dekker, and A. Broekers-Knol. 2020. "Schriftelijke Antwoorden Op Vragen Gesteld Tijdens de Eerste Termijn van de Begrotingsbehandeling van Het Ministerie van Justitie En Veiligheid (35570-VI)," November 25, 2020. [https://www.tweedekamer.nl/kamerstukken/brieven\\_regering/detail?id=2020Z22827&did=2020D48307](https://www.tweedekamer.nl/kamerstukken/brieven_regering/detail?id=2020Z22827&did=2020D48307).
- Gray, Jonathan, Carolin Gerlitz, and Liliana Bounegru. 2018. "Data Infrastructure Literacy." *Big Data & Society* 5 (2): 205395171878631. <https://doi.org/10.1177/2053951718786316>.
- Greenland, Sander. 2002. "Causality Theory for Policy Uses of Epidemiological Measures." In *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*, edited by Christopher J. L. Murray, Joshua A. Salomon, Colin D. Mathers, Alan D. Lopez, and Weltgesundheitsorganisation, 291–302. Geneva: World Health Organization.
- Griffin, Gabrielle, and Rosi Braidotti. 2001. "Whiteness and European Situatedness." In *Thinking Differently: A Reader in European Women's Studies*, edited by Gabrielle Griffin and Rosi Braidotti, 221–37. London: Zed Books.
- Groenendijk, Kees. 2007. "Allochtonen of burgers: definitiemacht in debat en wetgeving over immigranten." In *Macht en verantwoordelijkheid: essays voor Kees Schuyt*, edited by Jan Willem Duyvendak, G. Engbersen, and M. Teeuwen, 101–12. Amsterdam: Amsterdam University Press.
- Grosfoguel, Ramón. 2004. "Race and Ethnicity or Racialized Ethnicities?: Identities within Global Coloniality." *Ethnicities* 4 (3): 315–36. <https://doi.org/10.1177/1468796804045237>.
- Gunawan, Basuki, and O. D. van den Muijzenberg. 1967. "Verzuilingstendenties En Sociale Stratificatie in Indonesië." *Sociologische Gids* 14 (3): 146–58.

- Haan, Ido de. 2010. "Imperialism, Colonialism and Genocide. The Dutch Case for an International History of the Holocaust." *BMGN - Low Countries Historical Review* 125 (2–3): 301–27. <https://doi.org/10.18352/bmgn-lchr.7123>.
- Hacking, Ian. 1986. "Making up People." In *Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought*, edited by Thomas C. Heller, Morton Sosna, and David E. Wellberry, 161–71. Redwood City, CA: Stanford University Press. [http://www.icesi.edu.co/blogs/antro\\_conocimiento/files/2012/02/Hacking\\_making-up-people.pdf](http://www.icesi.edu.co/blogs/antro_conocimiento/files/2012/02/Hacking_making-up-people.pdf).
- Haggerty, Kevin D., and Richard V. Ericson. 2000. "The Surveillant Assemblage." *British Journal of Sociology* 51 (4): 605–22. <https://doi.org/10.1080/00071310020015280>.
- Hall, Stuart. 1992. "Cultural Studies and Its Theoretical Legacies." In *Cultural Studies*, edited by L. Grossberg, C. Nelson, P.A. Treichler, L. Baughman, and J.M. Wise. New York: Routledge.
- . 1995. "When Was 'the Post-Colonial'? Thinking at the Limit." In *The Postcolonial Question: Common Skies, Divided Horizons*, edited by Iain Chambers and Lidia Curti, 242–60. London: Routledge.
- , ed. 1997a. *Representation: Cultural Representations and Signifying Practices*. London, Thousand Oaks & New Delhi: SAGE Publications.
- . 1997b. "The Spectacle of the 'Other.'" In *Representation: Cultural Representations and Signifying Practices*, edited by Stuart Hall, 223–90. London, Thousand Oaks & New Delhi: SAGE Publications.
- Hall, Stuart, Chas Critcher, Tony Jefferson, John Clarke, and Brian Roberts. 1978. *Policing the Crisis: Mugging, the State and Law and Order*. Basingstoke, Hampshire: Palgrave Macmillan.
- Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575–99. <https://doi.org/10.2307/3178066>.
- . 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge.
- Harding, Sandra. 1991. *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca, NY: Cornell University Press.
- Hartigan, John. 1997. "Establishing the Fact of Whiteness." *American Anthropologist* 99 (3): 495–505.
- Hart-Kemper, R., and J. Nas. 1998. *Exit. Een Onderzoek Naar de Vertrekredenen van Expolitiemedewerkers van Het Regiokorps Amsterdam-Amstelland*. Amsterdam: Projectbureau Diversiteit en Emancipatie.
- Heide, Lars. 2009. *Punched-Card Systems and the Early Information Explosion, 1880–1945*. Baltimore, MD: Jon Hopkins University Press.
- Heidegger, Martin. 1977a. "The Question Concerning Technology." In *The Question Concerning Technology, and Other Essays*, translated by William Lovitt, 3–35. New York, NY: Garland Pub.
- . 1977b. *The Question Concerning Technology, and Other Essays*. Translated by William Lovitt. New York, NY: Garland Pub.
- Heijs, Eric. 1995. "Van Vreemdeling Tot Nederlander: De Verlening van Het Nederlandschap Aan Vreemdelingen 1813-1992." Dissertation, Nijmegen: Katholieke Universiteit Nijmegen.

- Helsloot, John I.A. 2012. "Zwarte Piet and Cultural Aphasia in the Netherlands." *Quotidian. Journal for the Study of Everyday Life* 3 (1): 1–20.
- Hiernaux, Jean. 1969. "Biological Aspects of the Racial Question." In *Four Statements on the Race Question*, edited by UNESCO, 9–16. Paris: Unesco. <http://unesdoc.unesco.org/images/0012/001229/122962eo.pdf>.
- Hochstenbach, Cody, Justus Uitermark, and Wouter van Gent. 2015. "Evaluatie Effecten Wet Bijzondere Maatregelen Grootstedelijke Problematiek ('Rotterdamwet') in Rotterdam." Amsterdam: University of Amsterdam.
- Hollinger, David A. 2005. *Postethnic America: Beyond Multiculturalism*. New York, NY: Basic Books.
- Homan, Gerlof D. 1966. "Catholic Emancipation in the Netherlands." *The Catholic Historical Review* 52 (2): 201–11.
- Hoogenboom, Hans. 2015. "De tapijtfabrikant en de dominee." *Hilversums Historisch Tijdschrift - Eigen Perk*, 2015.
- Hoogenboom, Marcel, and Peter Scholten. 2008. "Migranten en de erfenis van de verzuiling in Nederland: Een analyse van de invloed van de verzuiling op het Nederlandse migrantenbeleid (circa 1970-heden)." *Beleid en Maatschappij* 35 (2): 107–24.
- Hutchby, Ian. 2001. "Technologies, Texts and Affordances." *Sociology* 35 (2): 441–56.
- Ihde, Don. 2009. *Postphenomenology and Technoscience: The Peking University Lectures*. SUNY Series in the Philosophy of the Social Sciences. Albany, NY: SUNY Press.
- Iliadis, Andrew. 2013. "Informational Ontology: The Meaning of Gilbert Simondon's Concept of Individuation." *Communication+1* 2 (1): 1–21.
- Iliadis, Andrew, and Federica Russo. 2016. "Critical Data Studies: An Introduction." *Big Data & Society* 3 (2): 1–7. <https://doi.org/10.1177/2053951716674238>.
- Isin, Engin, and Evelyn Ruppert. 2019. "Data's Empire: Postcolonial Data Politics." In *Data Politics: Worlds, Subjects, Rights.*, edited by Didier Bigo, Engin Isin, and Evelyn Ruppert, 207–27. London & New York: Routledge.
- Jennissen, Roel, Frank Bovenkerk, and Godfried Engbersen. 2021. "Clusteren op basis van cultuur." Working Paper 48. Den Haag: WRR.
- Jennissen, Roel Peter Wilhelmina, Godfried Engbersen, Meike Bokhorst, and M. A. P Bovens. 2018. "De nieuwe verscheidenheid: toenemende diversiteit naar herkomst in Nederland." Den Haag: WRR.
- Jordan, Reed (Reed K. ). 2018. "The Dutch Urban Areas Act : A Black Liberation Perspective on the Policy of Exclusion." Thesis, Cambridge, MA: Massachusetts Institute of Technology. <https://dspace.mit.edu/handle/1721.1/118207>.
- Kam, George de, and Eelkje Idsardi. 2019. "Opvattingen van Stichting WAG-Deelnemers over de Effecten van Aardbevingen Op Het Woongenot En de Woningwaarde in Groningen." URSI Research Report 347. Groningen: Rijksuniversiteit Groningen. <https://core.ac.uk/download/pdf/232469147.pdf>.
- Keij, Ingeborg. 2000. "Standaarddefinitie Allochtonen." *Index [Statistics Netherlands Magazine]* 10 (November/December): 24–25.
- Kellenbach, F.M. 1999a. "Mechanisering En Automatisering." In *Een Eeuw Statistieken: Historisch-Methodologische Schetsen van de Nederlandse Officiële Statistieken in de*

- Twintigste Eeuw*, edited by B. Erwich and J.G.S.J. van Maarseveen, 71–104. Amsterdam: CBS.
- . 1999b. “Van Telmachine Naar Knooppunt Op de Elektronische Snelweg.” In *Welgeteld Een Eeuw*, edited by J.G.S.J. van Maarseveen and R. Schreijnders, 63–70. Amsterdam: CBS.
- Kennedy, Helen, and Rosemary Lucy Hill. 2018. “The Feeling of Numbers: Emotions in Everyday Engagements with Data and Their Visualisation.” *Sociology* 52 (4): 830–48. <https://doi.org/10.1177/0038038516674675>.
- Kennedy, Helen, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. 2016. “The Work That Visualisation Conventions Do.” *Information, Communication & Society* 19 (6): 715–35. <https://doi.org/10.1080/1369118X.2016.1153126>.
- Kitchin, Rob. 2014a. “Big Data, New Epistemologies and Paradigm Shifts.” *Big Data & Society* 1 (1): 1–12.
- . 2014b. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.
- . 2021. *Data Lives: How Data Are Made and Shape Our World*. Bristol: Bristol University Press.
- Kitchin, Rob, and Tracey P. Lauriault. 2018. “Toward Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work.” In *Thinking Big Data in Geography: New Regimes, New Research*, edited by Jim Thatcher, Andrew Shears, and Josef Eckert, 12–24. Nebraska: University of Nebraska Press. <https://papers.ssrn.com/abstract=2474112>.
- Kittler, Friedrich. 1990. *Discourse Networks 1800/1900*. Translated by Michael Metteer and Chris Cullens. Stanford, CA: Stanford University Press.
- . 1999. *Gramophone, Film, Typewriter*. Translated by Geoffrey Winthrop-Young and Michael Wutz. Stanford, CA: Stanford University Press.
- Kleijn, Koen. 2019. “De Groene Salon / The Green Room.” Edited by Gijs Frieling and Iris Thewessen. Rijksvastgoedbedrijf.
- Kool, Linda, Jelte Timmer, and Rinie van Est. 2015. “De Datagedreven Samenleving.” Den Haag: Rathenau Instituut.
- Kop, Nicolien, and Peter Klerks. 2009. *Doctrine Intelligencegestuurd Politiewerk*. Apeldoorn: Politieacademie. <https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/72506.pdf>.
- Koyuncu-Lorasdağı, Berrin. 2013. “‘Ethnicization of Islam’ and Headscarved Dutch-Turkish Students’ Identity Politics in the Netherlands: The Case of Amsterdam.” *Uluslararası İlişkiler / International Relations* 10 (38): 57–75.
- Kuijlaars, Anne-Marie. 1999. *Huis Der Getallen. De Institutionele Geschiedenis van Het Centraal Bureau Voor de Statistiek (CBS) En de Centrale Commissie Voor de Statistiek (CCS), 1899-1996*. Amsterdam: CBS.
- Kuiler, Erik W. 2014. “From Big Data to Knowledge: An Ontological Approach to Big Data Analytics.” *Review of Policy Research* 31 (4): 311–18. <https://doi.org/10.1111/ropr.12077>.
- Kunkeler, Nathaniël. 2018. “NARRATIVES OF DECLINE IN THE DUTCH NATIONAL SOCIALIST MOVEMENT, 1931–1945.” *The Historical Journal* 61 (1): 205–25. <https://doi.org/10.1017/S0018246X17000188>.

- Laan Bouma-Doff, Wenda van der. 2007. "Confined Contact: Residential Segregation and Ethnic Bridges in the Netherlands." *Urban Studies* 44 (5–6): 997–1017. <https://doi.org/10.1080/00420980701255965>.
- Latour, Bruno. 1991. "Technology Is Society Made Durable." In *A Sociology of Monsters: Essays on Power, Technology and Domination*, edited by John Law, 103–32. London: Routledge.
- . 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Leefbaarometer. 2019. "Help." Leefbaarometer. 2019. <https://leefbaarometer.nl/page/Help#wat>.
- . 2021a. "Wat is een grid?" Leefbaarometer. 2021. <https://leefbaarometer.nl/page/Help#grid>.
- . 2021b. "Wat is leefbaarheid?" Leefbaarometer. 2021. <https://leefbaarometer.nl/page/Help#wat>.
- Leeuwen, H.D. van. 1968. "Verandering en onderwijs in Suriname." *Sociologische Gids* 15 (6): 379–92.
- Leidelsemeijer, K., G. Marlet, J. van Iersel, C. van Woerkens, and H. van der Reijden. 2008. "De Leefbaarometer: Leefbaarheid in Nederlandse wijken en buurten gemeten en vergeleken rapportage instrumentontwikkeling." 95640. Utrecht: RIGO & Atlas voor Gemeenten.
- Leidelsemeijer, K., G. Marlet, R. Ponds, R. Schulenberg, and C. van Woerkens. 2015. "LEEFBAAROMETER 2.0: INSTRUMENTONTWIKKELING." Utrecht: RIGO & Atlas voor Gemeenten.
- Leidelsemeijer, K., and I. Van Kamp. 2003. "Kwaliteit van de Leefomgeving en Leefbaarheid: Naar een begrippenkader en conceptuele inkadering." Amsterdam and Bilthoven: RIGO & RIVM.
- Leun, J.P. van der, M. Kromhout, M. Easton, and F. Weerman. 2010. "Criminaliteit, Migratie En Etniciteit. Nieuwe Richtingen Binnen Een Complex En Beladen Onderzoeksterrein." *Tijdschrift Voor Criminologie* 52 (2): 107–21.
- Leun, J.P. van der, and M.A.H. van der Woude. 2011. "Ethnic Profiling in the Netherlands? A Reflection on Expanding Preventive Powers, Ethnic Profiling and a Changing Social and Political Context." *Policing and Society* 21 (4): 444–55. <https://doi.org/10.1080/10439463.2011.610194>.
- Leun, J.P. van der, M.A.H. van der Woude, R.D. Vijverberg, R.P.M. Vrijhoef, and A.J. Leupen. 2014. "Etnisch Profileren in Den Haag? Een Verkennend Onderzoek Naar Beslissingen En Opvattingen Op Straat." Leiden: Universiteit Leiden, Faculteit der Rechtsgeleerdheid, Afdeling Criminologie.
- Leurs, Koen, and Sandra Ponzanesi. 2018. "Connected Migrants: Encapsulation and Cosmopolitanization." *Popular Communication* 16 (1): 4–20. <https://doi.org/10.1080/15405702.2017.1418359>.
- Leurs, Koen, and Tamara Shepherd. 2017. "Data and Discrimination." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es. Amsterdam: Amsterdam University Press.

- Leurs, Koen, and Kevin Smets. 2018. "Five Questions for Digital Migration Studies: Learning From Digital Connectivity and Forced Migration In(to) Europe." *Social Media + Society* 4 (1): 2056305118764425. <https://doi.org/10.1177/2056305118764425>.
- Lewis, Heather A. 2015. "Math Mistakes That Make the News." *PRIMUS* 25 (2): 181–92. <https://doi.org/10.1080/10511970.2014.928657>.
- Lijphart, Arend. 2018. "Consociationalism After Half a Century." In *Consociationalism and Power-Sharing in Europe*, edited by Michaelina Jakala, Durukan Kuzu, and Matt Qvortrup, 1–9. Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-67098-0\\_1](https://doi.org/10.1007/978-3-319-67098-0_1).
- Linke, Uli. 1999. *Blood and Nation: The European Aesthetics of Race*. Philadelphia, PA: University of Pennsylvania Press.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Lovin, Clifford R. 1967. "Blut Und Boden: The Ideological Basis of the Nazi Agricultural Program." *Journal of the History of Ideas* 28 (2): 279. <https://doi.org/10.2307/2708423>.
- Lum, Cynthia, and William Isaac. 2016. "To Predict and Serve?" *Significance* 13 (5): 14–19. <https://doi.org/10.1111/j.1740-9713.2016.00960.x>.
- Lum, Cynthia, and Christopher S. Koper. 2015. "Evidence-Based Policing." In *Critical Issues in Policing: Contemporary Readings*, edited by Roger G. Dunham and Geoffrey P. Alpert, 260–74. Long Grove, IL: Waveland Press.
- Madianou, Mirca. 2019. "Technocolonialism: Digital Innovation and Data Practices in the Humanitarian Response to Refugee Crises." *Social Media + Society* 5 (3): 205630511986314. <https://doi.org/10.1177/2056305119863146>.
- Maldonado-Torres, Nelson. 2007. "On the Coloniality of Being: Contributions to the Development of a Concept." *Cultural Studies* 21 (2): 240–70. <https://doi.org/10.1080/09502380601162548>.
- Mali, Bas. 2020. "Predictive Policing in Nederland." *Cahiers Politiestudies* 54 (1): 89–102.
- Mali, Bas, Carla Bronkhorst-Giesen, and Mariëlle Den Hengst. 2017. *Predictive Policing, Lessen Voor de Toekomst*. Apeldoorn: Politieacademie.
- Mantello, Peter. 2016. "The Machine That Ate Bad People: The Ontopolitics of the Precrime Assemblage." *Big Data & Society* 3 (2).
- Marchand, Iris. 2014. "Dogla Politics? Questioning Ethnic Consociationalism in Suriname's National Elections of 25 May 2010." *Ethnic and Racial Studies* 37 (2): 342–62.
- Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York, NY: Houghton Mifflin Harcourt Publishing Company.
- M'charek, Amade. 2000. "Technologies of Population: Forensic DNA Testing Practices and the Making of Differences and Similarities." *Configurations* 8 (1): 121–58. <https://doi.org/10.1353/con.2000.0005>.
- . 2005. "The Mitochondrial Eve of Modern Genetics: Of Peoples and Genomes, or the Routinization of Race." *Science as Culture* 14 (2): 161–83. <https://doi.org/10.1080/09505430500110945>.
- . 2013. "Beyond Fact or Fiction: On the Materiality of Race in Practice." *Cultural Anthropology* 28 (3): 420–42. <https://doi.org/10.1111/cuan.12012>.

- M'charek, Amade, and Irene van Oorschot. 2020. "What about Race?" In *The Routledge Companion to Actor-Network Theory*, edited by Anders Blok, Ignacio Farías, and Celia Roberts, 235–45. Routledge Companions. Abingdon, Oxon New York, NY: Routledge.
- M'charek, Amade, Katharina Schramm, and David Skinner. 2014. "Technologies of Belonging: The Absent Presence of Race in Europe." *Science, Technology, & Human Values* 39 (4): 459–67. <https://doi.org/10.1177/0162243914531149>.
- McLeod, John. 2010. *Beginning Postcolonialism*. Second Edition. Manchester: Manchester University Press.
- McNeil, Maureen. 2005. "Introduction: Postcolonial Technoscience." *Science as Culture* 14 (2): 105–12. <https://doi.org/10.1080/09505430500110770>.
- McPherson, Tara. 2012. "Why Are the Digital Humanities so White? Or Thinking the Histories of Race and Computation." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 139–60. Minneapolis & London: University of Minnesota Press.
- Meesters, Paulien, Jan Kortekaas, and Maaïke Tragter. 1999. "Intelligence Led Policing: Nieuw Concept Voor Integratie van Oude Adagia." *Tijdschrift Voor Criminologie* 41 (4): 419–29.
- Meijer, A.H. 1992. *Straatnamenboek van Noordwijk: Noordwijks Historie Vanuit de Straatnaam*. Hilversum: Verloren.
- Meijer, Albert, Lukas Lorenz, and Martijn Wessels. 2021. "Algorithmization of Bureaucratic Organizations: Using a Practice Lens to Study How Context Shapes Predictive Policing Systems." *Public Administration Review* 81 (5): 837–46. <https://doi.org/10.1111/puar.13391>.
- Meijer, Albert, and Martijn Wessels. 2019. "Predictive Policing: Review of Benefits and Drawbacks." *International Journal of Public Administration*, February, 1–9. <https://doi.org/10.1080/01900692.2019.1575664>.
- Meijer, Remco. 2017. "Groot tekort aan agenten dreigt: komende jaren 14.000 politiemensen met pensioen." *de Volkskrant*. November 29, 2017. <https://www.volkskrant.nl/gs-bea9ff63>.
- Mepschen, Paul. 2012. "Gewone mensen. Populisme en het discours van verdringing in Amsterdam Nieuw West." *Sociologie* 8 (1): 66–83. <https://doi.org/10.5117/SOC2012.1.MEPS>.
- . 2016. "Everyday Autochthony: Difference, Discontent and the Politics of Home in Amsterdam." Dissertation, Amsterdam: University of Amsterdam.
- Meyrowitz, Joshua. 1999. "UNDERSTANDINGS OF MEDIA." *Review of General Semantics* 56 (1): 44–52.
- Milan, Stefania, and Emiliano Treré. 2019. "Big Data from the South(s): Beyond Data Universalism." *Television & New Media* 20 (4): 319–35. <https://doi.org/10.1177/1527476419837739>.
- Mills, Charles Wade. 1997. *The Racial Contract*. New York, N.Y.: Cornell University Press.
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. 2002. *The Constitution of the Kingdom of the Netherlands 2002*. The Hague: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

- Molenaar, Joseph, Paul Kroes, Carin Esman, Hani Quint, Bart Collard, Ricardo Lanza, and Jan Nap. 2020. "Professioneel Controleren - Een Onderzoek Naar Het Effect van de Pilot Proactief Controleren.Pdf." Apeldoorn: Politieacademie.
- Molina Espeleta, Natalia, Marieke Ruitenbergh, Myrthe Toppen, and Samra el Haji. 2021. "Verkleurde Beelden - Klachtbehandeling Etnisch Profileren." 2021/030. Den Haag: Nationale Ombudsman. [https://www.nationaleombudsman.nl/system/files/bijlage/Verkleurde%20Beelden-%20klachtbehandeling%20etnisch%20profileren%20DEF\\_1.pdf](https://www.nationaleombudsman.nl/system/files/bijlage/Verkleurde%20Beelden-%20klachtbehandeling%20etnisch%20profileren%20DEF_1.pdf).
- Mooij, Martine de, Wouter van Andel, Bart Bakker, Ruben van Gaalen, Rosa Meijer, Magda Sloodbeek, Mersiha Tepic, and Caroline van Weert. 2015. "Herziening Gewichtenregeling Voor Het Onderwijsachterstandenbeleid Primair Onderwijs." Den Haag: CBS.
- Morozov, Evgeny. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: Public Affairs.
- Muis, Iris, Gerwin van Schie, Maranke Wieringa, and Tim de Winkel. 2019. "'Liberation Begins with Stating the Facts': Framing Statistics and Information Bricolage in Geert Wilders' Twitter Practice." *Open Library of Humanities* 5 (1): 1. <https://doi.org/10.16995/olh.320>.
- Mulder, Hans. 1982. "Ontaard En Gezond. Nationaal-Socialistische Kunstcritiek in Nederland." In *Berlijn - Amsterdam 1920 - 1940: Wisselwerkingen*, edited by Kathinka Dittrich, Paul Blom, and Flip Bool, 314–23. Amsterdam: Querido.
- Mutsaers, Paul. 2014. "'All of Me': Psychologizing Turkish-Dutch Police Officers in the Netherlands." *Anthropology of Work Review* 35 (2): 72–83. <https://doi.org/10.1111/awr.12041>.
- . 2015. "A Public Anthropology of Policing: Law Enforcement and Migrants in the Netherlands." Dissertation, Tilburg: Tilburg University.
- Nakamura, Lisa. 2008. *Digitizing Race: Visual Cultures of the Internet*. Edited by Katherine Hayles, Mark Poster, and Samuel. Weber. Vol. 23. *Electronic Mediations*. Minneapolis, MN: University of Minnesota Press.
- . 2013. *Cybertypes: Race, Ethnicity, and Identity on the Internet*. London: Routledge.
- Nix, Justin. 2015. "Predictive Policing." In *Critical Issues in Policing: Contemporary Readings*, edited by Roger G. Dunham and Geoffrey P. Alpert, 275–88. Long Grove, IL: Waveland Press.
- Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Norman, Donald A. 1988. *The Design of Everyday Things*. New York, NY: Basic Books.
- . 1999. "Affordance, Conventions, and Design." *Interactions* 6 (3): 38–43.
- NOS. 2017. "Personeelstekort bij de politie: 'In veel plaatsen piept en kraakt het.'" NOS. 2017. <https://nos.nl/l/2245189>.
- . 2019. "Genetici helpen bij 'dna-muur' koning, gevoelige informatie weggelaten." nos.nl. April 7, 2019. <https://nos.nl/l/2292019>.
- NRC. 2009. "De Stad Leeft Op." NRC. December 19, 2009. <https://www.nrc.nl/nieuws/2009/12/19/de-stad-leeft-op-11827412-a161378>.
- Olsthoorn, Peter. 2016. "Big Data Voor Fraudebestrijding." Working Paper 21. Den Haag: WRR. <http://szwweb.adlibhosting.com/Details/..%5CPDF%5C72887.pdf>.



- Omi, Michael. 1997. "Racial Identity and the State: The Dilemmas of Classification." *Law & Inequality* 15 (1): 7–23.
- Omi, Michael, and Howard Winant. 2015. *Racial Formation in the United States*. Third Edition. Critical Social Thought. New York and London: Routledge & Kegan Paul.
- Onuoha, Mimi. 2018. "Notes on Algorithmic Violence." *Github.Com* (blog). 2018. <https://github.com/MimiOnuoha/On-Algorithmic-Violence>.
- Oosterloo, Serena, and Gerwin van Schie. 2018. "The Politics and Biases of the 'Crime Anticipation System' of the Dutch Police." In *Proceedings of the Workshop on Bias in Information, Algorithms and Systems*, edited by Jo Bates, Paul D. Clough, Robert Jäschke, and Jahna Otterbacher, 2103:30–41. Sheffield: CEUR Workshop Proceedings.
- Orlow, D. 2016. *The Lure of Fascism in Western Europe: German Nazis, Dutch and French Fascists, 1933-1939*. New York, NY: Palgrave Macmillan.
- Ostaijen, Mark van, and Peter Scholten. 2017. "The Politics of Numbers. Framing Intra-EU Migrants in the Netherlands." *Critical Policy Studies* 11 (4): 477–98. <https://doi.org/10.1080/19460171.2016.1224725>.
- Özdil, Zihni. 2014. "'Racism Is an American Problem': Dutch Exceptionalism and Its Politics of Denial." *Frame* 27 (2): 49–64.
- Perry, Walt L. 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Santa Monica, CA & Washington, DC: Rand Corporation.
- Peters, John Durham. 2015. *The Marvelous Clouds: Toward a Philosophy of Elemental Media*. Chicago and London: University of Chicago Press.
- Phan, Thao, and Scott Wark. 2021. "Racial Formations as Data Formations." *Big Data & Society* 8 (2): 20539517211046376. <https://doi.org/10.1177/20539517211046377>.
- Ploeg, Irma van der, and Isolde Sprenkels. 2011. "Migration and the Machine-Readable Body: Identification and Biometrics." In *Migration and the New Technological Borders of Europe*, edited by Huub Dijstelbloem and Albert Meijer, 68–104. Migration, Minorities and Citizenship. Houndmills, Basingstoke & Hampshire: Palgrave Macmillan.
- Politie Nederland. 2020. "Meetbare diversiteit politie is stabiel." June 22, 2020. <https://www.politie.nl/nieuws/2020/juni/22/meetbare-diversiteit-politie-is-stabiel.html>.
- Ponzanesi, Sandra. 2019. "Migration and Mobility in a Digital Age: (Re)Mapping Connectivity and Belonging." *Television & New Media* 20 (6): 547–57. <https://doi.org/10.1177/1527476419857687>.
- Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- . 2015. "The Flight of the Indicator." In *The World of Indicators*, edited by Richard Rottenburg, Sally Engle Merry, Sung-Joon Park, and Johanna Mugler, 34–55. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9781316091265.002>.
- Prianti, Desi Dwi. 2018. "Towards the Westernized Body: A Popular Narrative Reinforced by Men's Lifestyle Magazines in Indonesia." *Inter-Asia Cultural Studies* 19 (1): 103–16. <https://doi.org/10.1080/14649373.2018.1422351>.
- Prins, Baukje. 2000. *Voorbij de Onschuld: Het Debat over de Multiculturele Samenleving*. Amsterdam: Van Genneep.

- Prof. Hoxha. 2007. "GeenStijl: GS Presenteert: De Allochtoon-O-Meter." *Geenstijl* (blog). January 25, 2007. [https://www.geenstijl.nl/45923/gs\\_presenteert\\_de\\_allochtoonom/](https://www.geenstijl.nl/45923/gs_presenteert_de_allochtoonom/).
- Pugliese, J. 2007. "Biometrics, Infrastructural Whiteness, and the Racialized Zero Degree of Nonrepresentation." *Boundary 2* 34 (2): 105–33. <https://doi.org/10.1215/01903659-2007-005>.
- Puwar, Nirmal. 2004. *Space Invaders: Race, Gender and Bodies out of Place*. Oxford and New York: Berg.
- Rasmussen, Kim-Su. 2011. "Foucault's Genealogy of Racism." *Theory, Culture & Society* 28 (5): 34–51. <https://doi.org/10.1177/0263276411410448>.
- Rath, Jan. 1991. *Minorisering: De sociale constructie van "minderheden"*. Amsterdam: Sua.
- Research, Information and Statistics Amsterdam. 2021. "WIA Leefbaarheid 2003-2015." Tableau Public. 2021. <https://public.tableau.com/app/profile/ois.amsterdam/viz/WIAleefbaarheid2003-2015/Instructie>.
- Ricaurte, Paola. 2019. "Data Epistemologies, The Coloniality of Power, and Resistance." *Television & New Media* 20 (4): 350–65. <https://doi.org/10.1177/1527476419831640>.
- Rienks, Rutger, and Marc Schuilenburg. 2020. "Wat is er nieuw aan het voorspellen van criminaliteit?" *Cahiers Politiestudies* 2020–1 (54): 39–54.
- Rijksvastgoedbedrijf. 2019. "Kunst DNA Salon (Groene Salon)." Webpagina. [Rijksvastgoedbedrijf.nl](https://www.rijksvastgoedbedrijf.nl). Rijksvastgoedbedrijf. June 19, 2019. <https://www.rijksvastgoedbedrijf.nl/vastgoed/vastgoed-in-beheer/den-haag-paleis-huis-ten-bosch/in-beeld/kunst-dna-salon>.
- Risam, Roopika. 2019. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Evanston, IL: Northwestern University Press.
- Roth, Lorna. 2009. "Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity." *Canadian Journal of Communication* 34 (1): 111–36.
- Rottenburg, Richard, and Sally Engle Merry. 2015. "A World of Indicators: The Making of Governmental Knowledge through Quantification." In *The World of Indicators*, edited by Richard Rottenburg, Sally Engle Merry, Sung-Joon Park, and Johanna Mugler, 1–33. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9781316091265.001>.
- Ruijter, Ad de de. 1998. *Pluriform in uniform: diversiteit bij de politie Amsterdam-Amstelland, een overzicht van 1987-1998*. Lelystad: Stichting IVIO.
- Safransky, Sara. 2020. "Geographies of Algorithmic Violence: Redlining the Smart City." *International Journal of Urban and Regional Research* 44 (2): 200–218. <https://doi.org/10.1111/1468-2427.12833>.
- Said, Edward W. 1979. *Orientalism*. New York, NY: Random House, Inc.
- Saini, Angela. 2020. *Superior: The Return of Race Science*. London: 4th Estate Ltd.
- Salomon, Jean-Jacques. 1984. "What Is Technology? The Issue of Its Origins and Definitions." *History and Technology* 1 (2): 113–56. <https://doi.org/10.1080/07341518408581618>.
- Sánchez-Querubín, Natalia, and Richard Rogers. 2018. "Connected Routes: Migration Studies with Digital Devices and Platforms." *Social Media + Society* 4 (1): 1–13. <https://doi.org/10.1177/2056305118764427>.

- Scannell, R. Joshua. 2019. "This Is Not Minority Report: Predictive Policing and Population Racism." In *Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life*, edited by Ruha Benjamin, 107–29. Durham, NC: Duke University Press. <http://public.eblib.com/choice/publicfullrecord.aspx?p=5779780>.
- Schäfer, Mirko Tobias, and Karin van Es, eds. 2017. *The Datafied Society: Studying Culture Through Data*. Amsterdam: Amsterdam University Press.
- Schendel, Sascha van. 2019. "The Challenges of Risk Profiling Used by Law Enforcement: Examining the Cases of COMPAS and SyRI." In *Regulating New Technologies in Uncertain Times*, edited by Leonie Reins, 225–40. Information Technology and Law Series. The Hague: T.M.C. Asser Press. [https://doi.org/10.1007/978-94-6265-279-8\\_12](https://doi.org/10.1007/978-94-6265-279-8_12).
- Schie, Gerwin van. 2018. "Origins: A History of Race-Ethnic Categorisation in the Dutch Governmental Data Ontology (1899–2018)." *TMG - Journal for Media History* 21 (2): 67–88. <https://www.tmgonline.nl/articles/10.18146/2213-7653.2018.367/>.
- . forthcoming. "The Datafication of Racialization and the Pursuit of Equality: The Case of the 'Barometer Culturele Diversiteit.'" In *Situating Data: A Cultural Inquiry*, edited by Nana Verhoeff and Karin van Es, 89–102. Amsterdam: Amsterdam University Press.
- Schie, Gerwin van, 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 Kerezi Klára and Veronika Nagy, 169–96. The Hague: Eleven International.
- Schie, Gerwin van, Alex Smit, and Nicolás López Coombs. 2020. "Racing through the Dutch Governmental Data Assemblage: A Postcolonial Data Studies Approach." *Global Perspectives* 1 (1). <https://doi.org/10.1525/gp.2020.12779>.
- Schinkel, Willem. 2010. "From Zoëpolitics to Biopolitics: Citizenship and the Construction of 'Society.'" *European Journal of Social Theory* 13 (2): 155–72. <https://doi.org/10.1177/1368431010362300>.
- . 2013. "The Imagination of 'Society' in Measurements of Immigrant Integration." *Ethnic and Racial Studies* 36 (7): 1142–61. <https://doi.org/10.1080/01419870.2013.783709>.
- Schoping, F. van. 1941. *Wien Neêrlands Bloed....: Het Rassenvraagstuk En Zijn Beteekenis Voor Nederland*. Amsterdam: Volk en Bodem.
- Schreven, Luuk. 2004. "Documentatierapport Volkstelling 1971." SAV-102413. Den Haag: CBS.
- Schrover, Marlou. 2010. "Pillarization, Multiculturalism and Cultural Freezing. Dutch Migration History and the Enforcement of Essentialist Ideas." *BMGN - Low Countries Historical Review* 125 (2–3): 329–55. <https://doi.org/10.18352/bmgn-lchr.7124>.
- Schulte Nordholt, Eric, Marijke Hartgers, and Rita Gircour, eds. 2004. *The Dutch Virtual Census of 2001. Analysis and Methodology*. Voorburg/Heerlen: Statistics Netherlands.
- Segura, María Soledad, and Silvio Waisbord. 2019. "Between Data Capitalism and Data Citizenship." *Television & New Media* 20 (4): 412–19. <https://doi.org/10.1177/1527476419834519>.
- Sekuur, M. A. 2010. "Representaties van Vogelaarwijken." Master, Rijksuniversiteit Groningen. <https://frw.studenttheses.ub.rug.nl/1713/>.
- Shachar, Ayelet. 2009. *The Birthright Lottery: Citizenship and Global Inequality*. Cambridge, MA: Harvard University Press.

- Sharma, Sanjay. 2013. "Black Twitter? Racial Hashtags, Networks and Contagion." *New Formations* 78 (1): 46–64.
- Simondon, Gilbert. 2015. "Culture and Technics (1965)." Translated by Olivia Lucca Fraser. *Radical Philosophy* 189 (Jan/Feb): 17–23.
- . 2017. *On the Mode of Existence of Technical Objects*. Translated by Cécile Malaspina and John Rogove. Minneapolis, MN: Univocal Publishing.
- Snelders, Stephen. 2007. "OP WEG NAAR EEN 'GERMAANSCH' VOLKSGEZONDHEID. NATIONAAL- SOCIALISME, ERFELIJKHEIDSLEER EN EUGENETICA IN NEDERLAND 1940-1945." *Gewina* 30 (2007): 62–74.
- Spivak, Gayatri Chakravorty. 1988. "Can the Subaltern Speak?" In *Marxism and the Interpretation of Culture*, edited by Cory Nelson and Lawrence Grossberg, 271–313. Urbana & Chicago: University of Illinois Press.
- Spivak, Gayatri Chakravorty, and Sarah Harasym. 2014. *The Post-Colonial Critic: Interviews, Strategies, Dialogues*. New York & London: Routledge.
- Squires, Gregory D., and Frank Woodruff. 2019. "Redlining." In *The Wiley Blackwell Encyclopedia of Urban and Regional Studies*, edited by Anthony M Orum, 1st ed., 1–8. Wiley. <https://doi.org/10.1002/9781118568446.eurs0260>.
- Stănescu, Vasile. 2018. "'White Power Milk': Milk, Dietary Racism, and the 'Alt-Right.'" *Animal Studies Journal* 7 (2): 103–28.
- Steadman, Ian. 2013. "Big Data and the Death of the Theorist." *Wired*. 2013. <http://www.wired.co.uk/article/big-data-end-of-theory>.
- Steenbrink, Karel. 2007. *Catholics in Indonesia, 1808-1942: A Documented History*. Vol. 2: The Spectacular Growth of a Self Confident Minority, 1903–1942. Leiden: KITLV Press.
- Stoler, Ann Laura. 1989. "Rethinking Colonial Categories: European Communities and the Boundaries of Rule." *Comparative Studies in Society and History* 31 (1): 134–61. <https://doi.org/10.1017/S0010417500015693>.
- . 1995. *Race and the Education of Desire: Foucault's History of Sexuality and the Colonial Order of Things*. Durham, NC: Duke University Press. <https://doi.org/10.1215/9780822377719>.
- . 2002. "Colonial Archives and the Arts of Governance." *Archival Science* 2 (1–2): 87–109. <https://doi.org/10.1007/BF02435632>.
- . 2009. *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*. Princeton, NJ: Princeton University Press.
- . 2016. *Duress: Imperial Durabilities in Our Times*. Durham & London: Duke University Press.
- Strahilevitz, Lior Jacob. 2008a. "Privacy versus Antidiscrimination." *University of Chicago Law Review* 75 (1): 363–82.
- . 2008b. "Reputation Nation: Law in an Era of Ubiquitous Personal Information." *Northwestern University Law Review* 102 (4): 1667–1738.
- Strategische Beleidsgroep Intelligence. 2008. "Waakzaam Tussen Wijk En Wereld: Nationaal Intelligence Model Sturen Op En Met Informatie." Politieacademie. <https://www.politieacademie.nl/kennisonderzoek/kennis/mediatheek/PDF/69628.pdf>.

- Stroshine, Meghan S. 2015. "Technological Innovations in Policing." In *Critical Issues in Policing: Contemporary Readings*, edited by Roger G. Dunham and Geoffrey P. Alpert, 229–43. Long Grove, IL: Waveland Press.
- Tannen, Deborah, Heidi E. Hamilton, and Deborah Schiffrin, eds. 2015. *The Handbook of Discourse Analysis*. Second Edition. Malden, MA: Wiley-Blackwell.
- Taylor, Chloë. 2011. "Race and Racism in Foucault's Collège de France Lectures: Race and Racism." *Philosophy Compass* 6 (11): 746–56. <https://doi.org/10.1111/j.1747-9991.2011.00443.x>.
- Taylor, Linnet. 2021. "There Is an App for That: Technological Solutionism as COVID-19 Policy in the Global North." *The New Common*, March, 209–15. [https://doi.org/10.1007/978-3-030-65355-2\\_30](https://doi.org/10.1007/978-3-030-65355-2_30).
- Thurlings, J. M. G. 1979. "Pluralism and Assimilation in the Netherlands, with Special Reference to Dutch Catholicism." *International Journal of Comparative Sociology* 20 (1–2): 82–100. <https://doi.org/10.1163/156854279X00184>.
- Timan, Tjerk, and Francisca Grommé. 2019. "A Framework for Social Fairness – Insights From Two Algorithmic Decision – Making Controversies in the Netherlands." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3756755>.
- Todorov, Tzvetan. 1993. "Race and Racism." In *Theories of Race and Racism: A Reader*, edited by Les Back and John Solomos, 64–70. Routledge Student Readers. London: Macmillan Education UK. [https://doi.org/10.1007/978-1-349-22911-6\\_2](https://doi.org/10.1007/978-1-349-22911-6_2).
- Torre, E.J. van der, P.J. Gieling, M.C. Dozy, F.C. van Leeuwen, and W. Hamoen. 2011. *Veilig politiewerk: de basispolitie over geweldgebruik*. Den Haag: Boom Lemma.
- Tromp, Reinier, Rik Delhaas, and David Davidson. 2021. "In het vizier van het Rotterdamse bijstandsfraude-algoritme? - Argos." VPRO. December 18, 2021. <https://www.vpro.nl/argos/lees/onderwerpen/artikelen/2021/in-het-vizier-van-het-rotterdamse-algoritme-.html>.
- Uitermark, Justus, Cody Hochstenbach, and Wouter van Gent. 2017. "The Statistical Politics of Exceptional Territories." *Political Geography* 57 (March): 60–70. <https://doi.org/10.1016/j.polgeo.2016.11.011>.
- UNESCO. 1969. *Four Statements on the Race Question*. Paris: Unesco. <http://unesdoc.unesco.org/images/0012/001229/122962eo.pdf>.
- Van der Burg, M.J. 2007. "Nederland Onder Franse Invloed: Cultuurtransfer En Staatsvorming in de Napoleontische Tijd, 1799-1833." Dissertation, Amsterdam: University of Amsterdam.
- Van der Haar, Marleen, and Dvora Yanow. 2011. "Allochtoon Als Metafoor En Categorie. Over de Handelingsimplicaties van Beleidstaal." *Beleid En Maatschappij* 38: 160–78.
- Van Dijck, José. 2014. "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology." *Surveillance & Society* 12 (2): 197–208.
- Verbeek, Peter-Paul. 2012. "Expanding Mediation Theory." *Foundations of Science* 17 (4): 391–95. <https://doi.org/10.1007/s10699-011-9253-8>.
- Vernooij, Joop. 2017. "Lepra en Katholieke Kerk in Suriname (19e Eeuw): een korte historie over de relatie lepra en de Rooms-Katholieke kerk in Suriname." *Academic Journal of Suriname* 8: 762–70.
- Verweij-Jonker, Hilda. 1971. *Allochtonen in Nederland*. Den Haag: Staatsuitgeverij.

- Vliet, Diantha Y. 2020. "Postcolonial Pete: Race, Media, and Memory in the Politics of Dutch Identity." Philadelphia, PA: Temple University.
- Vossen, Koen. 2010. "Populism in the Netherlands after Fortuyn: Rita Verdonk and Geert Wilders Compared." *Perspectives on European Politics and Society* 11 (1): 22–38. <https://doi.org/10.1080/15705850903553521>.
- Vries, Katja de. 2020. "AI Policy in the Netherlands: More Focus on Practice than Principles When It Comes to Trustworthiness." In *Human-Centred AI in the EU Trustworthiness as a Strategic Priority in the European Member States*, edited by Stefan Larsson, Claire Ingram Bogusz, and Jonas Andersson Schwarz. Brussels: European Liberal Forum. [https://fores.se/wp-content/uploads/2021/02/Larsson\\_et\\_al\\_2020\\_AI-in-the-EU.pdf](https://fores.se/wp-content/uploads/2021/02/Larsson_et_al_2020_AI-in-the-EU.pdf).
- Vuijsje, Herman. 1986. *Vermoorde Onschuld: Etnisch Verschil Als Hollands Taboe*. Amsterdam: Bert Bakker.
- Vukov, Tamara. 2016. "Target Practice: The Algorithmics and Biopolitics of Race in Emerging Smart Border Practices and Technologies." *Transfers* 6 (1): 80–97. <https://doi.org/10.3167/TRANS.2016.060107>.
- Weiner, Melissa F. 2014a. "(E)RACING SLAVERY: Racial Neoliberalism, Social Forgetting, and Scientific Colonialism in Dutch Primary School History Textbooks <sup>1</sup>." *Du Bois Review: Social Science Research on Race* 11 (2): 329–51. <https://doi.org/10.1017/S1742058X14000149>.
- . 2014b. "The Ideologically Colonized Metropole: Dutch Racism and Racist Denial: Dutch Racism." *Sociology Compass* 8 (6): 731–44. <https://doi.org/10.1111/soc4.12163>.
- . 2015. "The Demography of Race and Ethnicity in The Netherlands: An Ambiguous History of Tolerance and Conflict." In *The International Handbook of the Demography of Race and Ethnicity*, edited by Rogelio Sáenz, David G. Embrick, and Néstor P. Rodríguez, 575–96. International Handbooks of Population. Dordrecht: Springer Netherlands. [https://doi.org/10.1007/978-90-481-8891-8\\_27](https://doi.org/10.1007/978-90-481-8891-8_27).
- Wekker, Gloria. 2016. *White Innocence: Paradoxes of Colonialism and Race*. Durham, NC: Duke University Press.
- Wekker, Gloria, and Helma Lutz. 2001. "A Windswept Plain. The History of Gender and Ethnicity Though in the Netherlands Translation of Een Hoogvlakte Met Koude Winden. De Geschiedenis van Het Gender En Etniciteitsdenken in Nederland." In *Caleidoscopische Visies. De Zwarte, Migranten En Vluchtelingen Vrouwenbeweging in Nederland*, edited by M Botman, Nancy Jouwe, and Gloria Wekker. Amsterdam: Koninklijk Instituut voor de Tropen.
- Welle Donker, F., R.C. Braggaar, and B. van Loenen. 2019. "Hergebruikers van Open Data in Beeld." Delft: TU Delft, Faculteit Bouwkunde.
- Wevers, Rosa. 2018. "Unmasking Biometrics' Biases: Facing Gender, Race, Class and Ability in Biometric Data Collection." *TMG - Journal for Media History* 21 (2): 89–105.
- Willems, Dick. 2017. "Predicting Criminal Events with Data Science." Presentation presented at the Data Science Utrecht Meet Up, Utrecht, October 24.
- Wimmer, Andreas, and Nina Glick Schiller. 2002. "Methodological Nationalism and beyond: Nation–State Building, Migration and the Social Sciences." *Global Networks* 2 (4): 301–34.
- . 2003. "Methodological Nationalism, the Social Sciences, and the Study of Migration: An Essay in Historical Epistemology." *International Migration Review* 37 (3): 576–610.

- Wohlers, Tony E., and Lynne Louise Bernier. 2015. *Setting Sail Into the Age of Digital Local Government: Trends and Best Practices*. New York: Springer.
- Woude, Maartje van der, and Joanne van der Leun. 2013. "De Nederlandse Veiligheidscultuur Als Katalysator Voor Etnisch Profileren?" *Tijdschrift over Cultuur En Criminaliteit* 2 (3): 123–36.
- WRR. 2016. *Big Data in een vrije en veilige samenleving*. Amsterdam: Amsterdam University Press.
- Yanow, Dvora. 2003. *Constructing "Race" and "Ethnicity" in America : Category-Making in Public Policy and Administration*. London and New York: Routledge.
- Yanow, Dvora, and Marleen Van der Haar. 2013. "People out of Place: Allochthony and Autochthony in the Netherlands' Identity Discourse—Metaphors and Categories in Action." *Journal of International Relations and Development* 16 (2): 227–61.
- 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>.
- Young, Robert J.C. 2004. *White Mythologies: Writing History and the West*, Second Edition. London & New York: Routledge.
- . 2012. "Postcolonial Remains." *New Literary History* 43 (1): 19–42.
- Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." SSRN Scholarly Paper ID 2594754. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2594754>.

## Summary in English

In the Netherlands, data about the spatial distribution of Dutch citizens structured on the basis of their “migration background” is available in governmental open data sets. In this dissertation, such data are called “race-ethnic” for the combined racial and ethnonationalist connotations of the categorizations and labels. Race-ethnic data are used by various governmental institutions to calculate crime risk, liveability scores for neighbourhoods, real estate value, welfare fraud risk, and many other variables that are deemed necessary in the governance of Dutch neighbourhoods and cities. I will argue that the appropriation of race-ethnic data for new purposes enables a process of racialization, understood as the selection of “particular human features for purposes of racial signification” (Omi and Winant 2015, 110). To understand how processes of racialization are not only historically and socially determined, but also result from specific technologies, I propose the concept of *technologically mediated racialization* (TMR).

TMR can be understood in an instrumental, epistemological, and ontological way. The instrumental approach to TMR focuses mainly on the content of a system. It invites questions about the definitions of race that are used, and about the type of datafied policy situation in which race is operationalized. The epistemological approach to TMR invites us to situate knowledge production through data and algorithms and show how this process is interwoven with culture, history, and ideology. The final, ontological approach to TMR asks questions about the relation of a data system to the world it is creating. Through this approach, we understand race in terms of a relational ontology that is dependent, not on biology or culture, but on the technologies that mediate its construction. In this understanding, race functions as a technicity; the situated racial formations produced by datafied technologies create specific social, technological, and cultural possibilities.

I use the instrumental, epistemological, and ontological approaches to TMR in the investigation of the development of past technologies used for counting and clustering people in the Dutch census between 1899 and 2011. I show how race-ethnic categorization is intimately tied to sociocultural ideas about which human characteristics matter, as well as to the available technologies to record, store, and distribute that information. I will trace the historical connection between contemporary Dutch categories and clusters, such as “person with a migration background”, and older categories and clusters, such as “niet-Westerse allochtoon”, and “vreemdeling”. This will allow me to show that contemporary categorization practices are still informed by Dutch colonial history and migration history. As a result, label changes are both unable to solve racializing practices, and, simultaneously, make historical connections between race-ethnic categories and colonial history harder to access. This process can be understood as a form of *colonial aphasia*: “a loss of access that may verge on active dissociation” (Stoler 2016, 12).



In my first two case studies, I show how race-ethnic categories are made available by the data infrastructure of Statistics Netherlands (CBS) and are used in data applications such as the Allochtoon-o-meter of the alt-right weblog *Geenstijl*, and the Leefbaarometer of the Dutch Ministry of the Interior. Where the Allochtoon-o-meter only provides percentages of people with a “Western” and “non-Western” migration background depending on the postal code it is given, the Leefbaarometer features a full map of the Netherlands that visualizes “liveability”, a measure that is partly based on the percentages of people with various migration backgrounds. While these two systems seem very different, at least in terms of their politics, these data applications inherit the same embedded perspective and normative assumptions about Dutchness already present in the data infrastructure of CBS. Therefore, the results of these applications are no accidental instances of TMR but rather materializations of a technicity of race that was already embedded in both the available technologies and related race-ethnic categories. Furthermore, especially in the case of the Leefbaarometer, the colonial and racialized connotations of the provided information have become obscured as a result of the decontextualizing nature of the process of datafication, worsening the aforementioned colonial aphasia.

My third case study, the Crime Anticipation System (CAS), a predictive policing system of the Dutch National Police, shows how even a system that does not use any explicit race-ethnic categories can produce TMR. In a *data assemblage* that includes both CAS, a system that aims to exploit information about societal inequalities to achieve operational “efficiency”, and an organizational culture that has only minimal protections against ethnic profiling, racialized and racializing results are no accident but an inherent feature of the system.

In the conclusions, I show how the order of the three approaches to TMR—from instrumental use, through epistemological critique, to ontological deconstruction—can be inverted and used in the planning, evaluation, and design processes of current and future governmental practices that are aided by data technologies. In this way, I hope that the results of this investigation are not only aiding in the production of knowledge within the scholarly fields of critical data studies and postcolonial studies, but also in guiding the Netherlands along a process towards a more accountable and just datafied society that values all people equally.

## Samenvatting in het Nederlands

De dataficatie van ras-ethniciteit: Een onderzoek naar technologisch gemedieerde racialisatie in datasystemen en -infrastructuren van de Nederlandse overheid

In Nederland zijn data over de geografische verdeling van Nederlandse burgers op basis van hun migratieachtergrond beschikbaar in Open Data van de overheid. In deze dissertatie worden dergelijke gegevens *ras-etnische* data genoemd vanwege de raciale en etno-nationalistische connotaties van de categorieën die worden gebruikt. Ras-etnische gegevens worden gebruikt door verschillende overheden voor het berekenen van risico's op criminaliteit, leefbaarheidsscores van wijken, huizenprijzen, risico's op bijstandsfraude, en vele andere zaken die belangrijk worden geacht voor het besturen van Nederlandse wijken, dorpen, en steden. Ik beargumenteer dat het gebruiken van ras-etnische data voor nieuwe doeleinden een proces van racialisatie in gang zet. Racialisatie kan worden begrepen als "de selectie van specifieke menselijke eigenschappen voor raciale betekenisgeving" (Omi en Winant 2015, 110; vertaling door de auteur). Om te begrijpen hoe processen van racialisatie niet alleen historisch en sociaal zijn gedetermineerd, maar ook het resultaat zijn van specifieke technologieën, stel ik het begrip *technologisch gemedieerde racialisatie* (TMR) voor.

TMR kan worden begrepen op een instrumentele, epistemologische, en ontologische manier. De instrumentele opvatting van TMR focust met name op de inhoud van een datasysteem. Met deze opvatting bevestig ik de gehanteerde definities van ras-ethniciteit, en de (gedataficeerde) beleidscontext waarin ras-etnische categorieën worden geoperationaliseerd. Met de epistemologische opvatting van TMR situeer ik de productie van kennis door data en algoritmes, waardoor de nadruk kan worden gelegd op de manieren waarop kennisproductie is verweven met cultuur, geschiedenis en ideologie. De ontologische opvatting van TMR stelt vragen over de relatie tussen een datasysteem en de wereld die het systeem helpt creëren. Met deze opvatting begrijpen we ras-ethniciteit als een relationele ontologie die niet afhankelijk is van biologie of cultuur, maar van de technologieën die de constructie van ras-ethniciteit mediëren. Hierdoor functioneert ras als techniciteit; gesitueerde raciale formaties die worden geproduceerd door datatechnologieën creëren specifieke sociale, technologische, en culturele praktijken en mogelijkheden voor ontwikkeling.

Ik gebruik de instrumentele, epistemologische en ontologische opvattingen van TMR in een onderzoek naar de ontwikkeling van technologieën die in het verleden werden gebruikt voor het tellen en categoriseren van mensen in de Nederlandse volkstellingen tussen 1899 en 2011. Ik laat zien dat ras-etnische categorisatie nauw verweven is zowel met socioculturele ideeën over welke menselijke eigenschappen Nederlandse overheden relevant achten, als met de technologieën die beschikbaar zijn voor de verzameling, opslag, en distributie van deze informatie. Ik bespreek de

connectie tussen huidige Nederlandse ras-etnische categorieën, zoals “persoon met een migratieachtergrond”, en oudere categorieën, zoals “niet-Westerse allochtoon” en “vreemdeling”. Dit maakt het mogelijk om te laten zien hoe huidige categorisatiepraktijken nog steeds beïnvloed worden door de Nederlandse koloniale geschiedenis en migratiegeschiedenis. Hierdoor zijn wijzigingen in vocabulaire, zoals de afschaffing van het woord “allochtoon”, niet in staat om racialiserende praktijken te voorkomen. Tegelijkertijd maakt de invoering van nieuwe labels het lastiger om te communiceren hoe ras-etnische categorieën gerelateerd zijn aan de Nederlandse koloniale geschiedenis. Dit proces kan worden begrepen als een vorm van *koloniale afasie*: “een verlies van toegang die neigt naar actieve dissociatie” (Stoler 2016, 12; vertaling door auteur).

In mijn eerste twee casestudies laat ik zien hoe ras-etnische categorieën beschikbaar worden gemaakt in de data-infrastructuur van het Centraal Bureau voor de Statistiek (CBS) en worden gebruikt in applicaties zoals de Allochtoon-o-meter van de alt-right website Geenstijl, en de Leefbaarometer van het Ministerie van Binnenlandse Zaken. Waar de Allochtoon-o-meter alleen percentages van mensen met een “Westerse” of “niet-Westerse” migratieachtergrond produceert, heeft de Leefbaarometer een volledige kaart van Nederland met een visualisatie van “leefbaarheid”, een indicator die deels wordt berekend op basis van de percentages van mensen van verschillende migratieachtergronden. Ondanks dat deze twee systemen zeer verschillende politieke doeleinden lijken te hebben, nemen ze beide het perspectief en de inherente normen van de data-infrastructuur van het CBS over. Daardoor kan de technologisch gemedieerde racialisatie, geproduceerd door beide applicaties, niet worden gezien als incidenteel, maar eerder als materialisatie van een techniciteit van ras-etniciteit die al aanwezig was in de beschikbare technologieën en ingebede categorieën. Daarbij worden met name in de Leefbaarometer de koloniale en geracialiseerde connotaties van de gebruikte data verdoezeld door de decontextualiserende werking van dataficatie, waardoor de eerdergenoemde koloniale afasie wordt versterkt.

Met mijn derde casestudie over het Criminaliteitsanticipatiesysteem (CAS) van de Nederlandse Politie laat ik zien hoe zelfs een systeem dat geen gebruik maakt van expliciete ras-etnische categorieën TMR kan produceren. In de *data assemblage* die bestaat uit zowel CAS, een systeem dat informatie over sociale ongelijkheid tracht te exploiteren om meer efficiëntie te bereiken, en een organisatiecultuur waarin er minimale bescherming is tegen etnisch profileren, komen geracialiseerde en racialiserende resultaten niet per ongeluk tot stand, maar dergelijke resultaten een inherent onderdeel van het ontwerp.

In de conclusie van mijn onderzoek laat ik zien hoe de volgorde van de drie opvattingen van TMR—van instrumenteel gebruik, via epistemologische kritiek, naar ontologische deconstructie—kan worden omgekeerd en worden gebruikt bij het plannen, evalueren, en

ontwerpen van huidige en toekomstige beleidsprojecten waarin data en algoritmes een rol spelen. Ik hoop dat de resultaten van dit onderzoek op die manier niet alleen bijdragen aan kennisproductie binnen de velden *critical data studies* en *postcolonial studies*, maar ook aan het begeleiden van Nederlandse overheden in het proces richting een eerlijke en transparante gedataficeerde samenleving waarin iedereen die in Nederland woont, gelijk wordt behandeld en gewaardeerd.

## Curriculum Vitae

Gerwin van Schie was born in Leiden, The Netherlands on January 2, 1987. He holds a BA degree in Theater, Film and Television Sciences and an MA degree in New Media and Digital Culture (Cum Laude), both from Utrecht University. Between 2015 and 2017, Gerwin worked at the Utrecht Data School as a Junior Researcher and at the Media and Culture Department of Utrecht University as a lecturer. In 2017, he started his NWO funded PhD project titled “Datafication of Race and Ethnicity in the Netherlands: Investigating Practices, Politics and Appropriation of Governmental Open Data” under the supervision of Sandra Ponzanesi and José van Dijck. During his PhD Gerwin published several articles about big data ethics, datafication, and race in the context of the Dutch government. On September 1st, 2022, Gerwin will start his new job as an Assistant Professor at the Media and Culture Department of Utrecht University.