Data-driven policing: Negotiating the legitimacy of the police

Fieke Jansen

Submitted for the degree of Doctor of Philosophy
Department of Journalism, Media and Culture
Cardiff University
January 2022

Abstract

The growing use of data-driven policing raises pertinent questions as to how the datafication of society changes the understanding of police power, crime, and justice. Media and surveillance scholars have examined the ideological grounds of datafication and the operations of global surveillance regimes. However, *how* data-driven policing tools are used and their impact on the justification and negotiation of police power have not been explored. To address these gaps this dissertation seeks to answer the following research questions: what is the nature of data-driven policing? And what is the relationship between datafication and police power? To answer these questions, I present original empirical research on two functions: data-driven risk scoring and biometric recognition in three countries and civic responses to their uses. I have engaged in fifty-six semi-structured interviews with experts, police practitioners, and civic actors; participant observation in police and civic meetings; and the study of grey literature.

My findings show that researching data as practice offers a nuanced account of to what end and on what grounds these tools become integrated within policing, and approaching data as struggle provides insights into what's at stake with the datafication of policing. I introduce the term *organisational optimisation logic* to foreground how data-driven policing is seen to 'fix' self-defined organisational challenges and allow the police to conform to the normative expectations datafication places upon public authorities. I put forward the term *politics of injustice* to account for the stratified ways data-driven policing invokes invisible hierarchies about whose voices count in the discussion around police power. These findings, I argue, have broader implications for how we come to understand police power and justice in a datafied society. I conclude by introducing the concept of *data legitimacy* to theorise on the relationship between datafication and the legitimacy claim of the police.

Table of Contents

Table of Figures

Acknowledgements

A number of people have been instrumental in the completion of this PhD. First and foremost, I would like to thank all my interviewees for taking the time to talk to me, share their thoughts, insights and experiences. I especially want to express my appreciation to the police practitioners who were willing to speak to an outside researcher and my gratitude to the social and racial justice actors for their openness and willingness to share their struggles with me. A huge thanks goes to my fantastic supervisor Lina Dencik, whose expertise, guidance, encouragement and patience has been invaluable throughout my PhD research, and Arne Hintz and Galina Miazhevich for their constructive feedback on my research. Needless to say, my understanding of data justice would not have been the same without the discussions with Lina Dencik, Philippa Metcalfe, Javier Sánchez Monedero, and Jędrzej Niklas, with whom I formed the Data Justice project. Thanks also to the Data Justice Lab and the Funding Matters group for showing that academia can be political and empowering.

There have been so many old and new friends who made the PhD writing process fun and gave life to the meaning of police power and justice. A few need special mention, Patrick Williams, Rebekah Delsol, Sarah Chander, Zara Rahman, and Nani Jansen Reventlow, our discussions on racial justice have been eye-opening, truly inspiring and at times uncomfortable. Nirit Peled and Jelle Klaas, our continuous collaboration allowed me to ground this research in my local context; that of inequality, repression, and risk scoring in Amsterdam. Niels ten Oever for being my friend and the only person crazy enough to volunteer to read the first draft of my first chapter and the first draft of my entire PhD. Again, thank you for the comments, support and distractions. Philippa Metcalfe, it was good to be on this ride together, our two week writing retreat on a Greek Island made finishing this PhD a lot more fun. Carly and Eric Kind, for the great dinners and conversations when I was in London, and your willingness to share contacts and ideas. Seeta Peña Gangadharan, and Sanne Stevens for continuously grounding these top-level conversations into what this mean for affected communities. Corinne Cath-Speth, for your shared critiquepatrick on safeguarding – a process to make repressive public systems seem less repressive. Michelle Thorne for your endless positivity and continuous encouragement. Finally, my family and friends for their support.

This PhD has been supported by the ERC Starting Grant DATAJUSTICE (grant no. 759903) under the Horizon 2020 research and innovation program.

1. Introduction

Artefacts of the increased datafication of society can be found across Europe, including in policing. In the United Kingdom, several police forces bought and tested commercial predictive policing tools that promise to predict where acquisitive crime is most likely to happen in the near future (Couchman, 2019; Jansen, 2018). In Belgium, the police turned to the testing of a facial-recognition system on the airport near Brussels after the terror attacks in 2016 (COC, 2019). In the Netherlands, the police built a predictive policing tool that aims to predict which adolescents and young adults are most at risk of engaging in criminal activities in the future (Abraham et al., 2011; Wientjes et al., 2017). In Brussels, the European Commission is investing in next-generation security technologies (Jones, 2017), and there is an ongoing discussion on expanding the nature and scope of Europol (Focant et al., 2012; Jones, 2011). While the introduction of many of these local and national projects have since been halted, these events demonstrate an increased interest in the development and testing of data-driven policing technologies for the purpose of crime control. I argue that this is indicative of a broader organisational transformation that is shaping the terms on which police power, crime, and justice are understood, which makes the relationship between datafication and the police an important site of study.

The introduction of predictive policing tools like PredPol, HunchLab, and Palantir in American police forces started critical debates on the impact of predictive analytics on the police and the policed. Prediction is seen to change the nature of policing from mostly reactionary, solving crime after it occurs, to more pre-emptive policing, deterring criminals by being at the right place at the right time (Brayne et al., 2015; Ferguson, 2017, 2012). Here, the promise of predictive analytics lies in its ability to optimise crime forecasts, which should allow the police to more efficiently allocate resources to pre-empt and reduce crime. However, scholars have shown that, as these tools build on historic police data to make predictions about the near future, they do not predict crime but rather forecast where police activity is most likely to happen (Van Brakel, 2016). A consequence of this approach is that, in a context in which police have historically over-policed communities of colour, it will predict future police actions in these same communities (Isaac and Lum, 2018; Lum and Isaac, 2016). Such accounts provide important insights into possible policing futures, which can materialise through the introduction of technology; however, it is imperative to recognise that there is still a lot unknown about the actual practice of data-driven policing (Brayne and Christin, 2021). The European policing context is even more under-researched; dominant scholarly and civic debates on data-driven policing started and centred around events in the United States. Therefore, I will argue that it remains unclear how data-driven policing becomes embedded within European

police forces and how their turn to data is changing the ways actors understand crime and just policing.

The perceived uptake of algorithmic modes of governance by European public institutions, including the police, is met with increased resistance from civil society. Through public reports, campaigns, and responses to European legislative debates, civil actors are raising human rights concerns, on topics such as privacy and anti-discrimination, and they are calling for safeguards and bans to mitigate potential harms that come from the datafication of the state (Access Now, 2021; EDRi, 2021; Niklas and Dencik, 2021). These critiques indicate that embedding forms of algorithmic governance (Katzenbach and Ulbricht, 2019) in public institutions is increasingly becoming a site for concern. The turn of the police, who are the symbol of state power and oppression, to data is considered to be particularly problematic. In recent years, civil society, primarily in the UK and in the Netherlands, has exposed and challenged a number of predictive policing (Amnesty International, 2020, 2018; Couchman, 2019) and facial-recognition pilots (Big Brother Watch, 2018; Liberty, 2018), arguing that they infringe on fundamental human rights and further entrench discriminatory practices within policing. These observations indicate that data-driven policing, as a medium through which European police forces engage with the world, is seen as a site of struggle that raises new concerns on what just and unjust policing should look like in a datafied society.

The police's turn to data-driven policing is not so much a new phenomenon but rather a continuation of police interest in data and technology for investigation and intelligence purposes (Bain, 2016). However, it is the introduction of more novel forms of data processing that are raising concerns about what their use means for how we come to understand crime and who is made visible to the police (Amoore, 2020; Taylor and Broeders, 2015). In this context, I build on scholarly debates that show that the construct of identity and risk are central to the police algorithmic gaze. Bureaucratic identity – such as last name, citizen identification number, and the integration of biometric data on passports – is a central component of modern statecraft (Scott, 1998). In contemporary societies, risk has become a central governance feature to mitigate the negative externalities of capitalism (Beck, 1992). The advent of predictive analytics is seen to further entrench identity and risk within governance practices, as the first allows those in power to tie data from seemingly disparate data infrastructures to a person (Gates, 2011; Jansen et al., 2021) and the latter allows historic data to predict present and future criminal behaviour (Harcourt, 2008). Yet how the datafication of society, and more specifically the perceived affordance of risk and

identification, are transforming policing practices related to crime prevention remains underexplored.

As alluded to so far in this introduction, the emergence of data-driven policing is both a policing practice and a site of struggle that is transforming how we come to understand crime and police power. Therefore, this thesis will approach the introduction of data-driven policing as a social process, which requires research to move beyond studying the technological artefact towards an inquiry into the social dimensions of its use. Here, I'm particularly interested in understanding how police power is perceived to become embodied and enacted through data systems. I will draw on social scientists, specifically Weber (1968), who have theorised that power in democratic societies is rarely enforced through the use of blunt force or the divine right to rule but has to be justified. As I will go on to argue in my thesis, police power enacted through data systems is productive and normative. I use the term productive to refer to the discourses, symbols, and norms around crime and justice that actively shape the image of the police and the state as a legitimate authority (Loader and Mulcahy, 2001) and create normative notions of who belongs or should be excluded from society (Williams and Clarke, 2018). Furthermore, the justification process is relational; it is a proposition made by the police that is responded to by different publics, which in some cases requires a response from the police (Bottoms and Tankebe, 2017). Thus, studying data-driven policing as a practice and a site of struggle will allow me to explore how its emergence is transforming the lines on which police power is justified and negotiated, and theorise about the relationship between datafication and police power.

In this thesis, I will present empirical research into the emergence of data-driven policing as a praxis and a site of struggle in the context of Europe. The central research questions are: what is the nature of data-driven policing? What is the relationship between datafication and police power? My analysis is based on fifty-six interviews with experts, police practitioners, and civic actors, substantiated with participant observation in police and civic meetings and the study of grey literature. Approaching data-driven policing as a practice will offer insights into the political rationale that underpins its introduction and brings nuance to the dominant understanding of to what end the constructs of risk and identity are deployed. In addition, engaging with data-driven policing as a site of struggle allows me to explore what is at stake when data as a medium mediates how we come to understand crime, policing, and justice. As I will argue, this approach allows me to situate the emergence of data-driven policing within the historic and ongoing process of the justification and negotiation of police power. Through this approach, I contribute to theory building on the relationship between datafication, police power, and justice.

A research approach into data-driven policing as a social process is like studying society: it is messy and complex, and it is characterised by different and at times conflicting interests, experiences, and discourses. To uncover and explain this 'mess', my thesis is rooted in critical realism and is empirically grounded. My data collection process maps what is happening in Europe, explores the actual practice of two data-driven policing functions, that of risk scoring and biometric recognition, and accounts for the responses to its emergence by listening to civic actors. Here, my research builds on Couldry's (2004) approach of media as practice that allows for studying data-driven policing as a sociotechnical system that is shaped by both the technology and its surroundings to offer insights into how these systems are used and perceived. I have operationalised my research through using the following methods: semi-structured interviews to listen to the experiences, attitudes, and beliefs of experts, police practitioners, and civic actors; participant observations in police and civic actor meetings; and the study of grey literature, such as reports, documentation, policy review, police presentations, and media reports.

1.1 Thesis outline

This dissertation consists of nine chapters, including this introduction. Like society, my research is contextual and composed of different layers. The central thread of my thesis is to reveal and explain the relationship between datafication and police power. I aim to account for the different forces that are at play in society by engaging in multi-sited empirical research to explore and connect perspectives from experts, police practitioners, civic actors, and external stakeholders. In chapter 2, I will outline how different fields of studies – such as media studies, surveillance studies, criminology, and social justice – inform my research, both their contributions and limitations. I will explore the dominant arguments on how datafication is perceived to transform the ontological and epistemological claims about knowledge, the contemporary economic paradigm, and modes of governance (boyd and Crawford, 2012; Cohen, 2019; Dencik et al., 2019; Lyon, 2006; Van Dijck, 2014). Here, I argue that media and surveillance study debates offer global theories on how power structures are reconfigured through the rise of complex and decentralised data infrastructures, yet provide little insight into how police make use of data and how this is changing the ways they come to understand and act upon crime. I will turn to debates in the field of criminology and social justice to explore frameworks on police power and justice, outside the context of datafication, to argue that, to fully account for what is changing with the introduction of data-driven policing, we need to situate it as part of a continuous negotiation process in which police power is enacted, justified, obeyed, and challenged.

Chapter 3 builds on the argument for studying data as practice by outlining the theoretical foundations and methodological choices that have informed my research approach. I start by discussing critical realism, a philosophy of science, which roots my research in the understanding that society is socially constructed, contextual, and stratified. Here, I explore the methodological benefits of engaging with data-driven policing as a social process to contribute to our understanding of the underlying structures and mechanisms that shape our society. Multi-sited empirical research into datafied policing, I argue, will offer insight into the political agenda and organising principles of data in relation to police practices and struggles on justice. Furthermore, it will allow me to connect these developments to the shifting social norms and values that shape its existence. Subsequently, I discuss my choice for studying data as practice, collecting data through case studies and semi-structured interviews, and substantiating these methods with the study of grey literature and participant observation; I will also discuss the limitations of these methods.

Chapter 4, my first empirical chapter, examines what is happening in the context of Europe by presenting the findings of my mapping study. Expanding on the discussion on data-driven policing, which are skewed either towards development in the US or focus on a specific technology or the legal regime that governs its use, my mapping will discuss the European policing context. Crucial in this mapping is the observation that the turn to data-driven policing is nascent and ephemeral in nature; the police are investing and experimenting with a broad range of tools that most likely will never materialise into every day of policing. However, my findings foreground a clear interest in and allocation of resources to specific data logics: augmenting databases, optimising operational support, real-time policing technologies, and predictive policing. In this chapter, I will also make the case for my case study approach, in which I will explore the function of data-driven risk scoring and biometric recognition across different jurisdictions.

Chapter 5, my first case study into data-driven risk-scoring projects, brings nuance to the dominant understanding of to what end risk is constructed within policing. Here, I engage with multi-sited research into four risk-scoring models, Top 600, ProKid, Integrated Offender Management model, and the domestic violence machine learning (ML) model, in the Netherlands and the UK. Through interviews, participant observation, and grey literature, I explore the political rationale from which these projects emerge, the meanings that become embedded within it, and the practices it invokes. To argue that, while dominant debates focus on how data-driven risk scoring attributes risk to an individual, practitioners primarily see these tools as a way ‘fix’ self-defined organisational challenges. To account for this observation, I will introduce the term *organisational optimising*

logic and argue that these findings have broader implications for how police power becomes embodied and enacted through data systems.

Chapter 6 expands on this discussion by presenting the findings of my second case study into biometric recognition. I will offer details on the development of facial-recognition and voice-identification systems in Belgium, Denmark, the European Commission (hereafter, Brussels), and the Netherlands. Their origin stories and practice offer a more nuanced understanding of to what end recognition becomes embodied and enacted within policing. This chapter therefore will examine the political rationale behind the use of facial recognition and voice identification to argue that the visible use of these technologies requires us to account for specific organisational aspirations. Where the turn to biometric recognition is seen to justify the police's desire to expand the nature and scope of policing and conform to the normative expectations of what contemporary public authorities should look like. Therefore, I will argue that biometric recognition speaks to a number of self-defined strategic and operational policing needs that again speak to the *organisational optimising logic* of data-driven policing.

Chapter 7 draws upon the study of civic responses to the emergence of data-driven policing to understand what is at stake, which allows me to explore the broader implications of datafication for how we come to understand police power and social justice. Through listening to civic actors involved in challenging police power, I will explore a range of injustice claims that relate to issues of discrimination, the criminal justice trap, data protection, governance, and access to justice. To account for what the introduction of data-driven policing means for our understanding of just and unjust policing, I will bring together these different concerns and explore how its use invokes invisible hierarchies on which voices count in the discussion on police power, what I call the *politics of injustice*. I will close the chapter by exploring what the politics of injustice means for the action strategies of civic actors, specifically the practice of reform and resistance.

Chapter 8 bring my findings on investment trends, police practices, and civic responses together to demonstrate their importance for theorising about the relationship between datafication and police power. This chapter outlines how the dominant frame of the managerial logic of datafication offers a limited understanding of how police power becomes embodied and enacted through data systems. I will start with an overview of my key findings, structured along my key concepts, the organisational optimisation logic and politics of injustice, to answer my first research question: what is the nature of data-driven policing? After which I will move towards my broader theoretical argument on the implications of the emergence of data-driven policing for our understanding of

police power and justice. I will build on social scientists Weber (1968), Beetham (1991a, 1991b), Bottoms and Tankabe (2012), and Martin and Bradford (2021) to explore the multitudinous and stratified ways in which the datafication of society becomes intertwined with questions of police legitimacy. I conclude by introducing the concept of *data legitimacy* to theorize on the dialogic relationship between datafication and the negotiation of police power.

Chapter 9 will conclude this dissertation by engaging with the theoretical and policy implications of my research. Here, I will explore how my research contributes to the scholarly debates on the relationship between datafication and power, and what my more nuanced account of what is at stake means for policy discussion and civic actions.

2. Debates on data, police power, and injustice

The datafication of society has given rise to a new field of research that explores how the increased volumes of data about people, objects, and events is transforming historically constructed power relations in society. Dominant scholarly debates from the disciplines of media and surveillance studies have immensely contributed to the understanding of how datafication is transforming ontological and epistemological claims about knowledge (boyd and Crawford, 2012; Van Dijck, 2014), the contemporary economic paradigm (Cohen, 2019; Zuboff, 2015), and power dynamics (Dencik et al., 2019; Lyon, 2006a). Data processing, as a dominant medium through which we come to organise and experience the world, is seen to invoke new modes of governance that are based on pre-emption and prediction (Andrejevic, 2020; Peters, 2013). These debates have informed this thesis on how data as a medium has come to shape the operations of power in contemporary society. However, as I will argue, connecting these insights to the operations of police power exposes two distinct knowledge gaps. Theories on data and power are global and abstract in nature, and as Dencik observed, 'we still struggle to account for the ways in which different actors make use of data, and how data is changing the ways actors understand and act in relation to social and political issues' (2019, p. 243). As such, if we assume, as this thesis does, that data as a medium is becoming integrated in policing, this in itself requires a deeper and contextual understanding of how it emerges as a practice and what this means for the perceptions of police power. The scholarly debates discussed in this chapter and the identified research gaps informed my research questions: what is the nature of data-driven policing? And what is the relationship between datafication and police power?

Centring power in my research on data-driven policing requires an articulation of how power is exercised, who is seen as the legitimate authority to wield power in a specific time and place, and why people consent to their demands. In my thesis, I draw on the broader field of social sciences, specifically Weber (1968), who theorises power as being relational, coercive, normative, and productive in nature. Police, criminologists argue, are continuously negotiating their role in society and justifying their actions, and they 'are deeply implicated in this production of the legitimate political orders which reproduces recognition of the state's right to assume this control' (Jackson and Bradford, 2009, p. 497). Thus, the police, through threat of punishment, play a key role in maintaining and reinforcing a vision of how society ought to look (Bourdieu, 1991; Loader, 1997). Police power, I will argue, is not static but rather political and productive in nature. The discourses, symbols, and norms around crime and justice actively shape the image of the police and the state as a legitimate authority (Loader and Mulcahy, 2001) and create normative notions of who belongs or

should be excluded from society (Williams and Clarke, 2018). Connecting these insights to the discussions within media and surveillance studies exposes a gap in the understanding of the relationship between datafication and the productive and relational nature of police power.

Any exploration of the relationship between data and police power requires me to bridge multiple scholarly fields. As such, I have chosen to do a topical literature review, and this chapter will outline how different fields of study have informed my research, both in its contributions and limitations. First, I start with exploring the scholarly debates that engage with the emergence of data-driven policing. Here, I found that existing research reveals that police are turning to contemporary data systems for their operations, yet there still is much unknown about this actual practice in Europe. Second, I zoom out and discuss the scholarly contributions from the field of media studies that theorise about what is changing now that data is becoming a prominent medium through which we order and control society. I will argue that these debates offer ample insights into the ideological grounds of data, but are less informative on how it shapes the understanding of crime, police power, and justice. Third, I turn to surveillance studies to explore debates on how the increased volume of data allow those in power to monitor and influence the present and future behaviour of people. I conclude that these theories contribute to knowledge about surveillance as a global mode of governance through which power holders are able to manage and control risk in society. Yet these theories offer limited insights into the productive nature of police power and decontextualise data-driven policing as a practice and a site for struggle. Finally, I turn to the social sciences, specifically debates in criminology and social justice, to explore theoretical debates that foreground how police power needs to be claimed, maintained, and legitimised, and how it invokes struggles over justice. To conclude, in exploring these interconnected but distinct disciplines, it becomes clear that there is still a lot unknown about the actual nature of data-driven policing and the relationship between datafication and police power.

2.1 Data-driven policing

The use of data and new technologies by police has been conceptualised as ‘predictive policing’, ‘intelligence-led policing’, and ‘algorithmic policing’. Each describes key aspects of the use of data by police: predictive policing emphasises its probabilistic and pre-emptive ability (Van Brakel, 2016); intelligence-led policing connects the increased interest in surveillance tactics to the shift from more traditional police practices to intelligence activities (Brayne, 2017); and algorithmic policing highlights the reliance on technology to make sense of the data for policing purposes (Egbert, 2019). What these concepts have in common is that they shed light on the emergence of

contemporary data systems as a new feature of policing and theorise about its effect on police perception of crime. In this section, I will explore the debates, predominantly from media and surveillance studies, that focus on the emergence of data and technology in the context of policing. I will refer to this as data-driven policing to encompass the multitude of technologies the police are interested in (Jansen, 2018). I find that, in particular, empirical studies into specific data-driven functions offer valuable insights into the interest and values that becomes embedded within policing through its use. However, I draw on Brayne and Christin (2021) and Flensburg and Lomborg (2021) to argue that there is still a lot unknown about the actual practice of data-driven policing, especially in Europe, and what its emergence means for our understanding of police power.

Debates on the police use of data and technology find their roots in the actuarial logic of prediction (Harcourt, 2008) and foreground how its introduction is shifting police operations from more traditional reactionary interventions, solving crime after it has occurred, to more proactive interventions, preventing crime from occurring or responding to it in real time (Brayne et al., 2015; Brayne, 2017; Dencik et al., 2018a; Van Brakel, 2016). In the broad range of technologies that are available to the police, I will elaborate on the discussion related to predictive policing systems, as these are currently the most debated data-driven policing function in media and surveillance studies. There are two distinct types of predictive policing systems: hotspot policing and predictive identification. Hotspot policing relies on identifying patterns in the distribution of crime to predict the location of where crime is most likely to happen in the near future (Kaufmann et al., 2019). This function is mostly directed at predicting high-impact crime, such as robbery, burglary, and theft, and can inform the extent to which police patrol certain areas. Predictive identification, also known as risk scoring, aims to predict who is most likely to become a potential offender or potential victim of a predefined crime priority. In most cases, police use these tools to identify, rank, and intervene in the lives of individuals who are already known to the police through a care and control approach (Ferguson, 2017; Van Schendel, 2019). The rationale for why police turn to predictive policing is primarily discussed in relation to the managerial logic of increased efficiency and effectiveness, in which statistical generated crime insights are believed to allow for better allocation of resources (Brayne, 2017; Ferguson, 2019, 2017; Hardyns and Rummens, 2018; Van Brakel, 2016).

The origins of the predictive policing debates stem from the use of hotspot policing in the United States, where scholars have explored data sets, algorithms, and data systems to theorise about its social implications (Ferguson, 2012; Hendrix et al., 2018; Pearsall, 2010; Perry et al., 2013; Shapiro, 2017). Researchers have foregrounded the use of two distinct models in hotspot policing. One model solely relies on police data – on the type of crime, location, and time – to predict where

high-impact crime is most likely to happen in the near future. The other model combines these data points with external variables – like weather, holidays, events, and distance to highways – to make its predictions (Ferguson, 2017; Hardyns and Rummens, 2018). While we can debate which model creates more accurate predictions, the central concern is that both models primarily rely on police data for prediction. Police databases are considered to be of poor quality, which leads to unreliable outputs (Kilkenny and Robinson, 2018), but more importantly, prediction on the basis of police data does not inform about crime; it informs about police activity. It will not predict where crime will most likely take place but where arrests are most likely to happen (Van Brakel, 2016). Lum and Isaacs (2018, 2016) re-creation of the predictive policing tool PredPol shows exactly this. They argue that prediction on the basis of historic police data creates a vision of the future that is based on the stereotypes of the past (Crawford, 2018), where the over-representation of Black[1] and Brown communities in police data will skew the prediction to these same communities. Thus, if and when the police will act on these results, it is argued to invoke a negative feedback of over-policing of already over-policed communities.

Empirical research into the actual practice of hotspot policing in the US and Europe offer a different understanding of what is changing with its introduction. Brayne's (2017) research into the use of big data by the Los Angeles Police Department shows how its existence invokes police interests in and needs for expanding data collection and surveillance efforts and creates a desire to make their existing databases interoperable. Similarly, in Europe, research shows that, despite disappointing results, the use of predictive policing systems is reinforcing the police's belief in and desire to work with data (Egbert, 2019; Egbert and Leese, 2021). Egbert found that 'it has made police authorities aware that the massive amounts of crime data they possess are quite valuable and can now be easily analyzed' (Egbert, 2019, p. 83), pointing to the organisational benefits of predictive systems that go beyond their intended purpose, in that they both reinforce a belief in the ability of data to better analyse and represent criminal activity and normalise the practice of data collection and algorithmic governance within the institution of policing. Furthermore, research into the actual use showed how the output of these models 'need not necessarily be "true" but merely accurate enough to inform operational measures' (Egbert and Leese, 2021, p. 3). The term 'operational measures' refers to adjusting the patrol frequency in neighbourhoods that are identified as being at risk. Researching the actual practice of predictive policing thus shows that these tools do not operate in isolation but become embedded within a broader organisational structure and culture.

[1]In this thesis, I will capitalize Black, Brown, and White when it refers socially constructed racial and ethnic identities.

The use of facial recognition in the context of policing is slowly becoming another area of scholarly interest. Where some researchers approach it as a technical system, others engage with it as a mode of governance that needs to be regulated, or as a distinct policing practice. Buolamwini and Gerbu (2018) have researched facial recognition as a technical artefact, showing how these recognition systems are error-prone and perform less well on certain demographics, especially Black women. These findings inform the public understanding that, when flawed systems become embedded within policing, they maintain and perpetuate the over-policing of communities of colour (Big Brother Watch, 2018). Kindt (2020) argues that the current European regulatory frameworks, such as the European General Data Protection Regulation (GDPR) and the Law Enforcement Directive (LED), are defunct, as they contain too many loopholes and broadly defined exceptions for the collection of biometric data. These scholars have significantly contributed to our understanding of the technical and legal challenges that arise with the introduction of facial recognition but say little about how these tools land in an operational organisation like the police. Fussey and Murrey (2019) conducted an independent, external evaluation on the London Metropolitan Police Service (the Met) trial on facial recognition, a deployment that has since been halted. They found that, in a live setting, police officers, or boots on the ground, rely on technology with the 'presumption to intervene', in which they overestimate the credibility of the system and will act more quickly upon its outcome. Connecting the technical limitations to the findings of the evaluation supports the broader concern that, when flawed systems are deployed in real time, they can perpetuate the over-policing of communities of colour.

I will conclude this section by observing that, while data collection and technology have always played a role in policing (Bain, 2016), these debates inform my research that data-driven policing is becoming a new feature of contemporary policing, and its introduction is accompanied by significant challenges. Empirical research into the actual practices shows that, beyond its intended purposes of identification and prediction, the use of data systems further reinforce the operational interest in and need for increased data collection and data-processing practices. Thus, data systems as such influence how the police come to define problems and solutions around crime and prevention. To fully account for what this shift towards data-driven policing means, the next section of this chapter will engage with media scholars who have theorised about datafication and the changing nature of power in contemporary societies. Beyond the insights into the unintended organisational changes that emerge from the introduction of hotspot policing and facial recognition, these empirical studies also offer a specific research approach. Engaging with the practice of data, what Couldry (2004) has called studying 'media as practice', allows this research to engage with the emergence of data-driven policing as a sociotechnical system that is shaped by both the technology

and its surroundings. In a context in which there is still a lot unknown about the actual practice of data-driven policing (Brayne and Christin, 2021), I argue that engaging with data as practice is particularly pertinent to explore how datafication is shaping the understanding of crime, police power, and justice.

2.2 Datafication and power

In this thesis, I understand the concept of power as the ability of an actor to manage or control the actions of others despite their resistance. The belief of an actor's legitimate authority and ability to justify its actions is a means to attain and maintain power in society (Weber, 1968; Beetham, 1991b). Power as such is not a static phenomenon but rather continuously challenged and negotiated, and paradigm shifts represent key moments in time in which power relations are transformed. Consequently, the increased datafication of society has given rise to an expanding body of scholarly work that engages with data as a medium through which power holders order and control society. Media scholars have positioned datafication as the latest paradigm shift that is significantly impacting the ontological and epistemological claims about knowledge (boyd and Crawford, 2012; Van Dijck, 2014), the contemporary economic paradigm (Cohen, 2019; Zuboff, 2015), and changing social structures (Dencik et al., 2019; Lyon, 2006a). Therefore, to understand the relationship between data and police power, as I intend to do in this thesis, I will now engage with the dominant debates primarily within the field of media studies to gain insight into the ideological grounds of datafication. This allows me to explore the shifts in what is considered relevant knowledge for governance and who is seen as the authority in creating said knowledge. To conclude that data-driven decision-making is increasingly becoming embedded within the social fabric of society and influencing how practitioners and civic actors interact with the world around them. However, these debates are less informative to understand how power becomes enacted through data systems.

The much-cited concept of 'datafication' by Mayer-Schoenberger and Cukier (2013) explains how increased integration of technologies in our everyday lives has allowed for human behaviour to be captured in data points that can be extracted, collected, and analysed for real-time tracking and predictive analytics. The datafication of society as such is seen to create an environment where data points get extracted and combined to create a digital representation about people, their behaviours, and their relationships to one another for economic or political purposes. The act of collecting, analysing, and using this data is what some media scholars see as the rise of logical media, an organisational medium that sorts people, objects, and events across time and space (Peters, 2013, p.

40). It is precisely this shift towards data as a means to interpret, order, and construct society that requires an articulation of what it means to see society through the quantified gaze. Van Dijck (2014) put forward the notion of ‘dataism’ to conceptualise that the transformation of data as a dominant form of knowledge production is rooted in a fundamental belief in its capacity to objectively represent social life and better predict individual behaviour than pre-datafication. This gives rise to an ideology that is rooted in the belief that data can ‘objectively quantify and potential tracking of all kinds of human behaviour and sociality through online media technologies’ (Van Dijck, 2014, p. 198). A belief that is dependent on a deep trust in both a causal relationship between data and people and in those entities that can process large amounts of data to predict human behaviour.

In his book *The Cultural Logic of Computation*, Golumbia (2009) argues that the belief in computation is not a new phenomenon but rather a continuation of the philosophical view of rationalism. For hundreds of years, Western societies have privileged logical reasoning as a way to manage society. Logic is positioned as value-ridden, objective, and neutral, and it holds authority in explaining and solving the world problems (Golumbia, 2009, pp. 189–191). Computation, as such, is not a radical break from old belief systems, but rather the emergence of data processing as a key mode of knowledge production is the result of the dominant world view that has shaped contemporary societies (Golumbia, 2009, p. 3). The meaning inscribed onto computation and data, to more objectively represent social life, is justified by the shared belief in logic and rationalism and, in effect, contributes to the preservation of historically determined power structures. Thus, when data processing becomes part of governance, it invokes normative claims on the neutrality, objectivity, and superiority of its decisions, which privileges a mode of governance that Andrejevic describes as ‘a persistent attempt to collapse the political into the technical as if the solution to societal and political conflicts were simply a matter of imperfect information’ (Andrejevic, 2020, p. 101). This observation connects the ideological grounds of computation to the desire to depoliticise the exercise and concentration of power, reducing social conflict to something that can be best managed through better data practices.

We can draw on Heidegger (2010) to understand the dialectic relationship between power and technology. He argued that technology is not just a means to an end, ‘techne’ creates its own interpretation of reality. It shapes the way we understand ourselves and the world around us (Hanks, 2010). In other words, when the police approach individuals, communities, human behaviour, and spaces through data, these become objects that can be ordered and arranged as raw material for production, which, according to Heidegger (2010), would inscribe specific meaning into how the

police come to understand problems, solutions, and the object of analysis. This understanding is one in which a society is no longer a complex organism of individuals, history, culture, and social norms, but where these become cluster of objects that can be organised and processed for a specific political or economic purpose (Ansorge, 2016, p. 116; Dencik, 2019; Waddington, 2005). This insight requires an articulation of what is changing in the ontological and epistemological claims of knowledge and the creation of subjects. Data-driven policing allows police to break away from historic knowledge production processes about individuals and society that were based on observation and human interpretation to a computational analysis that looks at the emergence of patterns about an individual, communities, and social relationships (Boyd and Crawford, 2011; Leese, 2020). An important characteristic of this type of knowledge production, finding patterns in vast amounts of data created and collected over time, is that it privileges correlation over causation, creating a paradigm in which the understanding of the 'what' is more important than the 'why' (Andrejevic, 2014). In his latest book, Andrejevic (2020, pp. 79–80) poses the example of predictive policing to argue that correlation as a dominant form of knowledge production creates a reality in which pre-emption is connected to the managing and controlling of social problems in the immediate rather than addressing its root causes – the unequal distribution of power, material resources, and life chances. Thus, power enacted through data systems is directed at managing unwanted behaviour over addressing the root cause of crime.

The datafication of society has further given rise to scholarly debates that engage with the changing institutional structures that produce knowledge from data. These discussions point to changes in who is seen as the authority in a datafied world. The material impact of the rise of big data, as argued by Boyd and Crawford (2011), is in how we come to think about 'the how' and 'the who' of knowledge production. Not only does it privilege machine-generated insights over other forms of knowledge production, but the centralisation of data infrastructures in the hands of commercial actors also raises significant questions over who has access to data and who has the ability to generate insights that inform our understanding of the world. Thus, the emergence of data as the raw material for knowledge production shifts the authority of sense-making to those actors who have access to the data, infrastructure, and means of computation (Boyd and Crawford, 2011; Hardy, 2014; Sadowski, 2019; Srnicek, 2017). What Andrejevic (2013; 2014) has labelled as the 'big data divide', in which the boundaries between those who create data, those who collect it, and those who analyse it, are being redrawn. These shifting boundaries of what and who is seen as the authority of knowledge and who has the means to produce it are marked by an increased asymmetry, what Harvey (2005, 2007), and Tatcher et al. (2016) refer to as the processes of 'accumulation by dispossession'. These developments have informed the political economy debate

on data and is argued to have given rise to a new capitalist class that owns and controls how information is gathered and used (Wark, 2021).

Critical political economy offers an approach to explore the relationship between data and capitalism. Among its scholars, there is general agreement that data is changing how the economy is organised, but there is much debate on the exact relationship between data and capitalism. Barbrook and Cameron's (1996) concept of the Silicon Valley ideology challenges the frame of the technology industry being captured by capitalism and exposes how the libertarian market logic was embedded in it from the start. Zuboff's (2015) much-cited concept of surveillance capitalism centres around 'behavioural surplus', in which she highlights how the data economy is centred around collecting and processing behavioural data, a by-product of people's technology use, to predict and modify 'human behavior as a means to produce revenue and market control' (Zuboff, 2015, p. 75). Cohen (2019) explores how legal systems are deeply implicated in the rise of the political economy of information. Srnicek's (2017) concept of platform capitalism explains that a competitive advantage in contemporary data economy requires companies to dominate the niche they operate in. For this, they need to continuously finding new ways of extracting data from individuals and objects and enhance their computing power to analyse this data. Sadowski (2019) argues that data is not just a commodity for production, but should be treated as capital, where the collection and circulation of data is not merely a means to an end but an end in itself. Wark's (2021) analysis of political economy returns to how power has shifted from the landlord class (who own the land) under feudalism to the capitalist class (who own the means of production), to a new kind of class that owns and controls how information is gathered and used.

Any top-level articulation of the political economy of data does not in any way do justice to the breadth and depth of scholarly work on this topic. Still, I choose to briefly acknowledge these debates and take from them the insight that datafication has given rise to a new economic paradigm that is centring data processing as a core element of value creation oriented towards continual growth (Thatcher et al., 2016). Here, it is important to note that data is perceived to be subjected to a different economy, one of abundance rather than scarcity, where, to gain and maintain a competitive advantage, companies are continuously finding new ways of extracting data from individuals and objects, and they are investing in the computing power to analyse this data (Brynjolfsson and McAfee, 2014; Srnicek, 2017). This has given rise to a new capitalist class that owns and controls how primarily consumer data is collected and processed. While it is unclear to what extent these developments directly impact European public institutions, and particularly European police forces, these market dynamics will have broader implications for their operations –

if only that these developments further inscribe authority to data to better interpret, order, and manage society than other forms of knowledge production and continuously present the police with new data collection possibilities and technical opportunities. These top-level insights are most relevant for my research, as the implications of the perceived value of data and the continuous need to redraw the boundaries on what can be made calculable and visible go well beyond the market, as I will show in the next section on surveillance and power.

These dominant debates within the field of media studies have significantly contributed to this thesis's understanding of the ideological grounds of data. The fundamental belief in data as an authoritative form of knowledge production as such can be seen to create normative claims of how society should be organised and who is best equipped to do organise it. This, I argue, highlights that attributes of power, normative and productive, are also ascribed to data. However, these ideological grounds offer little insight into how data systems manifest themselves in local contexts nor into how state power becomes enacted through data systems. To explore the latter, I will move on to the debates within the field of surveillance studies that theorise about the nature and operation of power in contemporary societies.

2.3 Power embodied and enacted through surveillance

A subset of scholarly critiques on data as a technology of governance stems from the field of surveillance studies. These studies concern themselves with the question of why and how people are tracked, surveilled, and governed (Lyon, 2001, 2007; Browne, 2015). Surveillance is conceptualised as an act of governance that is performed by the state and its institutions, the market and its companies, and society and its people. Scholars acknowledge that the act of surveillance, watching and being watched, long pre-dates the datafication of society. However, it is the rise of computation and its vastly expanding data infrastructures that have marked a turning point in what is believed to be possible in the classification and organisation of people, objects, and events across time and space (Ball et al., 2012; Haggerty et al., 2011; Lyon, 2006b). In this section, I will explore the dominant scholarly debates that foreground surveillance as a general and global social phenomenon, which has significantly contributed to this thesis's understanding of the relationship between data and state power. Surveillance is positioned as a central organising principle, which is enacted to ensure a specific vision of society and responded to by different publics; as such, power is not one-directional but rather dialectic and normative in nature. I will conclude this section by arguing that, in theorising about surveillance as an abstract global phenomenon, scholars have under-theorised

about its productive nature and at times missed the nuance of how it becomes integrated as a practice and a site of struggle within a specific policing context.

The Foucauldian (1977) use of Beetham's prison panopticon has been very influential within surveillance studies. The panopticon offers Foucault a blueprint to theorise about the shifting nature and operation of power in contemporary societies, where the power holder no longer controls its populations through naked repression or threats of physical violence, but through the internalised disciplinary gaze as a mechanism of governance (Elmer, 2012; Haggerty, 2006). Power thus operates through a process in which individuals are aware that they are subjected to a constant but unverifiable disciplinary gaze within a demarcated space, like the prison or the hospital, which results in self-regulation to fit the normative notion of what a good prisoner, patient, or citizen should look like. The disciplinary gaze as such is seen to invoke self-regulation, compliance, and submission of the individual and the masses, thus it is an enactment of power that allows for both reform of individual subjectivity and a mode of governance within industrial capitalism to manage different populations (Elmer, 2012; Haggerty, 2006; Ponterotto, 2016). This more subtle enactment of power, over that of physical punishment and blunt force, is assumed to modify individual and collective behaviour and produce docile subjects of governance (Elmer, 2012). However, empirical research in 'panoptic' prisons shows how strict implementation of the disciplinary gaze results, not in docile subjects, but rather in distinct forms of refusal and resistance, and makes it a site of struggle (Lyon, 2006b).

Foucault's use of the panopticon has equally inspired surveillance scholars to theorise about being watched as a means to invoke and coerce compliance of the individual and the masses, as it has been critiqued for its inadequacies to account for today's data economies (Bigo, 2006; Gandy Jr, 1993; Haggerty et al., 2011; Lyon, 2007; Wacquant, 2001). As a concept, the panopticon has informed numerous scholars theorising about different aspects of the relationship between data and power. Take, for example, Poster (1990), who put forward the concept of the 'superpanopticon' in 1990 to account for new modes of surveillance that are based on the electronic language of computing and allowed for a new kind of sense-making about people across time and space. Gandy Jr's (2021, 1993) proposition of a 'panoptic sort' already drew attention to the relationship between information and power as a means for economic production in 1993. He refers to this development as the political economy of personal data and argues that it has allowed for the intensification of classification and categorisation of people for commercial interest. Bigo's (2008) concept of the ban-opticon connects the post-September 11 security context with surveillance studies to argue that the state anti-terrorism discourse normalised increased monitoring and security practices on the

entire population for the purpose of excluding and controlling the dangerous and undesirable. These different interpretations of the panopticon try to account for the expanding spatio-temporal dimensions of surveillance beyond the boundaries of institutions and point to how the visible use of surveillance technologies and security politics are shifting social norms, in which being watched becomes normalised.

The first decade of the 2000s marked a theoretical shift away from the Foucauldian notion of the panopticon (Boyne, 2000; Lyon, 2006a). Surveillance scholars felt that the panopticon model was too limited for understanding the multifaceted ways in which power is established and enacted within complex data infrastructures. They argue that the datafication of society allowed for a new kind of decentralised monitoring of consumers and citizens, referred to as 'dataveillance' (Clarke, 1988; Raley, 2013), which is characterised by increased data creation and collection, often without a specific purpose in mind (Andrejevic, 2012). The changing modes and practices of surveillance, in which data processing have become embedded within the very fabric of society, was argued to require new ways of thinking about the operations of power (Lyon, 2006b). Here, it was not the intention to replace the panoptic model with another all-encompassing theory of surveillance; rather, influential scholars (Boyne, 2000; Haggerty, 2006; Lyon, 2006a) felt that a theoretical shift was needed to create room for a multiplicity of theories that could account for distinct aspects of surveillance.

This shift in surveillance studies has been characterised by a plurality of approaches, concepts, and case studies. Scholars increasingly theorise about the impact of internet infrastructure, from the rise of platform politics (Langlois and Elmer, 2019; Van Dijck et al., 2019) to the emergence of smart city technologies (Melgaço and Brakel, 2021), on the configuration of power relations in society. Others situate surveillance within a specific context, from policing (Brayne, 2017; Egbert, 2019) to the workplace (Ball, 2010; Dencik and Stevens, 2021), enacted within a specific geography (Bigo and Guild, 2005; Metcalfe and Dencik, 2019) or on specific communities (Browne, 2015). More recently, in response to the Covid-19 pandemic, scholars have started to explore what is at stake now that public health monitoring is connected to the everyday practice of monitoring and controlling people's behaviour (Newell, 2021), what Taylor et al. (2021, p. 11) have called the 'epidemiological turn in digital surveillance'. Again, others explored the function and limits of contemporary legal frameworks and oversight mechanisms (Kindt, 2020; Van Brakel, 2021a). In this plurality of approaches and theory building, I will engage with the concepts of control, identity, and recognition that emerged from the post-panoptic discussions, as these are most relevant for my research into the relationship between datafication and police power.

The concept of control emerged as a prominent post-panoptic understanding of surveillance and has primarily been informed by Deleuze's (2017) 'Postscript on the Societies of Control' (Elmer, 2012; Galič, Timan and Koops, 2017), in which power is no longer exercised through the internalised disciplinary gaze, but through a mode of governance, enabled by vast expanding and decentralised sites of surveillance, that uses data processing to control and better manage future behaviour (Brusseau, 2020). Since its publication, some surveillance scholars (Elmer, 2012; Galič, Timan and Koops, 2017) have argued that Deleuze's 'Postscript on the Societies of Control' (1992) is no longer an abstract concept, but an economic and political reality, in which the omnipresence of contemporary data infrastructures offer the possibility for the continuous and defuse monitoring of people, their movements, behaviours, and interests (Cohen, 2019; Zuboff, 2015). Control is, in the economic sense, enacted through predictive analytics to regulate behaviour through information filters, positive incentives, and nudging (Brusseau, 2020, p. 2). Coercion through incentives, might, at first sight, seem less relevant in the context of policing, as they are most known to enforce compliance through the threat of punishment. However, Henman's (2011) concept of 'new conditionalities' foregrounds how investments in the interoperability of databases allow public authorities to invoke compliance by making public services dependent on each other. For example, predictive identification programmes that promote a care and control approach place risk and security at the centre of governance, using both the city's physical infrastructure and public health and care authorities as environments through which criminality can be managed (Amnesty International, 2018; Ferguson, 2017; Haggerty et al., 2011; Jansen, forthcoming). Compliance is rewarded with access to other public services, and refusal and resistance is punished through disciplinary measures by both the police and the broader state.

Underneath the surface of the control society lie two concepts, identity and risk, that are closely tied to the rise and operations of predictive analytics. We can draw on Scott's (1998) analysis of the function of the state to understand how the construct of identity has become embedded within control as a mode of governance. Scott's description of the rise of the early modern nation state showed how standardisation, and in its wake bureaucratisation, allowed for a central authority to rule an unruly society. Here, legibility, tying an administrative identity to a person, became a central instrument of governance for the purposes of taxation and security, making the bureaucratic identity an intrinsic part of the state's acquisition, exercise, and maintenance of power. Thus, early forms of data collection did not merely strengthen the position of the nation state by increasing its finances, efficiency to rule, and ability to control its population by eliminating elements of surprise and risk, it further ingrained the affordance of national identity systems as a central instrument to advance the

administrative ordering of society (Brusseau, 2020; Haggerty et al., 2011). Since these early ledgers, advances in identity systems have allowed those watching to aggregate and assemble data from various data infrastructures, a practice that has become even more prominent in the state's attempts to keep up with the increased global mobility of both goods and people and has allowed for the enactment of the securitisation politics that followed the end of the Cold War (Gates, 2011; Jansen et al., 2021).

Identity systems are therefore important governance tools for monitoring both who enters and exits sovereign territories and how individuals behave domestically (Gates, 2011). Lyon (2008) has referred to the growing use of and dependency on identity systems as 'governing by identity', whereby legitimate status in society, access to basic services, and public and private space is increasingly tied to the ability to produce and verify someone's identity. In this sense, identity systems are perceived as a means by which the state can more effectively and efficiently engage in statecraft, border control, policing; administer public services; and put conditions on citizenship (Lyon, 2008; Van Zoonen, 2013). The state interest in and reliance on identity has propelled continuous technological developments that are directed at more reliably tying an administrative identity to a person. The latest shift is towards the use of biometrics systems, where the use of a digital representation of a fingerprint, face, or voice is believed to construct a reliable single identity that cannot change over time and is less susceptible to exploitation and abuse than other forms of identification (Gates, 2011, p. 14; Leese, 2020). Identity systems are therefore still central to governance and are slowly moving away from being organised around a unique identification number or a first and last name to being organised around bodily characteristics (Kak, 2020; Van Zoonen, 2013).

Beck's (1992) work on the 'risk society' explores how risk has become a central feature of governance in contemporary societies. He argues that, since the Industrial Revolution, society has entered in a process of 'reflexive modernisation', as man-made risks, an unwanted side effect of modernity, create social uncertainties that force societies to change. 'As a result of this process, society in the "second modernity" is no longer concerned with the distribution of power and wealth, but instead with the way it handles risks' (Wimmer and Quandt, 2007, p.337). Risk as a technology of governance can take two forms: the first fundamentally seeks to address the root causes of man-made risks, while the second focuses on managing the consequences of risks. The actuarial logic that underpins the personalisation of risk that informs predictive analytics (Harcourt, 2008) falls firmly in the second category. It uses abstraction and calculation to predict and control individuals that display concerning behaviour in relation to a predetermined social problem (Dencik, 2021;

Harcourt, 2008) instead of investing in and changing the root causes of risk. These debates contribute to our understanding of the values that become inscribed within data systems and is best described by this quote from Andrejevic:

> If, in the industrial era, the promise of automation was to displace manual labour, in the information age it is to pre-empt agency, spontaneity, and risk: to map out possible futures before they happen to objectionable ones can be foreclosed and the desirable ones selected. (Andrejevic, 2020, p. 9)

The debates on control, identity, and risk inform my thesis by foregrounding how the normative and coercive nature of power becomes embodied and enacted through data systems. Predictive analytics can be perceived as the key site of social conflict, as they invoke modes of governance that privilege the managing of the immediate over structurally intervening in the conditions that create social uncertainty.

I will close this section on surveillance by engaging with the scholarly debates that theorise about the ‘who’ of surveillance, which offers insight into the intent or the politics behind the act, behind data processes that monitor and categorise people. Amoore (2020) argues that we should engage with algorithmic systems as ‘ethicopolitical’ arrangements and scrutinise them for how they transform who and what is made visible and calculable. This foregrounds how surveillance assemblages (Haggerty and Ericson, 2000) are not neutral constructs that enact control; rather, they are dependent on who and what is defined as a target of interest for economic and political gain and thus becomes the object of surveillance. This process is described by Gandy Jr (1993) and Lyon (2003) as surveillance as social sorting, in which those who wield the power of surveillance can use it to verify both the identity of a person and assess their risk or worth. Here, it is the intent behind social sorting through which ‘people may be categorized and treated differently as a result of gender, race, ethnicity, religion, class and age, among other forms of difference’ (Lyon, 2007, p. 17). Fourcade and Healy (2017) theorise on how data as a probabilistic mechanism of governance contributes to the accumulation of symbolic and intangible capital that in turn shape an agent’s social, cultural, and economic position in society. An individual’s value or worth, in this sense, is construed through the meaning they are afforded by algorithmic classification systems (Fourcade and Healy, 2017). These scholarly debates all foreground how the introduction of data-driven decision-making in state services, including the police, is accompanied by significant challenges and is argue to further institutionalise the classification of ‘deserving’ and ‘undeserving’ citizens (Redden, 2018; Romano, 2017).

The dominant surveillance discussion about the who of surveillance privilege the act of surveillance as the primary organising principle through which the state engages with the world. I draw on Browne's (2015) work, which demonstrates that race as an organising principle of society is often overlooked within surveillance studies. In her book, Browne (2015) questions what would happen if surveillance theories did not start from the panoptic model, but with the branding and monitoring of Black and Brown communities during transatlantic slavery and its afterlife. She builds on critical race scholars such as Fanon (1968), Collins (1990), and Fiske (1998) to argue that categorising on the basis of race, class, and gender is not the result of but the raison d'être of surveillance as social sorting. Brown (2015) builds on Fiske's (1998) observation that Whiteness is the norm in contemporary societies and everything else, for example desirable and undesirable behaviour, is measured against it. Thus, the classification of deserving and undeserving citizens is not an unintended consequence of governance but a heavily racialised process that requires an understanding of how the White gaze of the state is historically determined and, to this day, ascribes meanings onto Black and Brown bodies (Benjamin, 2019; Hamilton, 2020; Omi and Winant, 1986). Browne's (2015) act of centring race within an analysis of surveillance as such exposes how racialised social structures are at the heart of historic and ongoing struggles for how society is organised and shows the need for these to be taken into account in an analysis of the relationship between datafication and police power.

I will draw on Garland's (2001) 'criminology of the other' to connect the organising principles of race and surveillance to that of police power. He situates the politicised discourse of crime, in which some criminals are portrayed as intrinsically evil or wicked and nothing like 'normal' upstanding citizens, within the political rationale of control. Stereotypes of the 'gang member', 'terrorist', and 'gipsy' position certain criminal acts as so immoral or such a fundamental threat to our collective safety that they constitute a group of perpetrators who cannot be rehabilitated but merely controlled. It is through the discourse surrounding the 'criminology of the other' that power holders construct a supposed link between race and crime. Williams and Clarke (2018) build on Garland's notion of the 'criminology of the other' to critique the construction of 'gang crime' in the context of the UK (Amnesty International, 2018; Williams, 2018; Scott 2018). They argue that the criminalising intent of this approach 'legitimizes intrusive racist policing and surveillance' (Williams and Clarke, 2018, p. 1). Thus, the productive power of the 'criminalization of the other' is that it further inscribed racialised notions of crime and legitimises increased surveillance activities that are directed at controlling some individuals and communities. Brown (2015), Garland (2004), and Williams and Clarke (2018) all show that, when we decentre the dominant understanding of surveillance as

control, a more nuanced understanding emerges on the coercive, normative, and productive nature of police power.

In this section, I have explored theories of surveillance that contribute to our understanding of the nature and modes of power that become embodied and enacted through data systems. These debates offer insights into the workings of surveillance as a general and global social phenomenon. Scholars have theorised on how the datafication of society allows for a new kind of monitoring and control in which data processing and prediction are used to regulate and manage present and future behaviour of people. These processes are dependent on the construct of identity and risk to tie data to a persona and assign worth to an individual and group. Specifically, Brown (2015) informs my thesis that centring data in an analysis of power run the risk of displacing other organising principles of society, such as race, class, and gender, that shape systems of control. Her observations require any research into data-driven policing to account for how surveillance and predictive analytics are fundamentally intertwined with the normative White gaze of the state that aims to control and manage those bodies and behaviours that do not conform. Furthermore, similar to the debates on the ideological grounds of data, these discussions offer limited insights into the knowledge gap identified by Dencik (2019) and Brayne and Christin (2021) that much is a lot unclear about how data-driven policing is used by police and how datafication is redrawing the lines on how we come to understand just and unjust policing. Therefore, I will now turn to debates in criminology and social justice that have informed my thinking on the operations of police power as productive and as a site of struggle.

2.4 The dialogic nature of power

The exercise and negotiation of power has been and remains a key debate in the social sciences. Thus far, this literature review has engaged with dominant debates from media and surveillance studies that contribute to my understanding of how power becomes embodied and enacted through data systems. I have argued that these debates offer insights into the impact of datafication on the coercive and normative nature of power enacted through global surveillance systems, but found them less informative about the nature of data-driven policing and how datafication is changing how we come to understand crime, police power, and justice. In this section, I will draw on social science debates that engage with questions of power, not data, which will offer me frameworks to understand the social dimensions of police power. I will start with Weber (1968), whose argument that, in democratic societies, state power is rarely enforced through the use of blunt force or justified on the divine right to rule but needs to be justified, has given rise to a number of legitimacy

debates. These insights inform this thesis's premise that police power is productive, as it depends on the belief in the police's legitimate authority and ability to justify its actions, and that police power is also relational, a dialogue between the operation of policing and the extent to which the publics find these actions just or unjust (Weber, 1968; Beetham, 1991b).

In his book *Economy and Society*, published posthumously, Weber (1968) observes that, in democratic societies, state power is rarely enforced through the use of blunt force or justified on the divine right to rule, but rather any formal system or organisation that aims to influence the behaviour of those who are subjected to it can only do so if they are believed to be legitimate (Beetham, 1991a; Bottoms and Tankebe, 2017). Thus, power can only be justified if both the power holder and those subjected to power believe it to be legitimate (Lee and McGovern, 2013). Here, legitimacy as a social construct operates in relation to the state, domination, and power (Wæraas, 2009). The state is defined as a compulsory political organisation that is tied to a specific territory in which they have the monopoly of violence and have been successful in upholding their claim to power. Domination is seen as a process of successfully imposing one's will on others, and power is conceptualised as 'an actor position to carry out his own will despite resistance' (Weber, 1968, p. 53). In his theory of power, Weber's (1968) term 'voluntary obedience' foregrounds that compliance to power is not a given, and resistance impacts the state's ability to rule and stay in power (Wæraas, 2009). This points to the relational nature of power, where the power holder enacts power from the belief that they have the moral right to govern (Tankebe, 2014), which is dependent on the extent to which individuals belief in their authority, obey the law, or comply with police actions. In this chapter, and in my research, I will approach them as two separate phenomena that are mutually dependent, to be able to understand why police turn to data and how publics respond to its emergence and the relationship between them.

Although Weber has significantly influenced the social and political science debate on legitimacy, his argument has equally been criticised. I want to highlight the main critique offered by Beetham (1991a), who observed that the relational nature of power is informed by and dependent on the dominant social norms and values that shape our societies. He argues that the power holder's legitimacy claim cannot solely be explained by the public belief in its legitimacy and instead requires a deeper understanding of other relevant factors that influence its perceived authority. It is not about the belief in legitimacy but 'power is legitimate to the extent that the rules of power can be justified in terms of beliefs shared by both dominant and subordinate' (Beetham, 1991b, p. xx). This slight nuance in the relationship between the power holder and the public positions it not only in relation to each other but also to the broader environment they are part of. For example, the

police self-belief that they have a rightful claim to power and enact certain modes of governance is reaffirmed by the deep-seated belief that the state is a superior institution that has knowledge others don't have to make fair decisions (Ansorge, 2016; Easterling, 2011). Furthermore, Beetham's observation can explain why people who do not believe in the legitimacy of police power feel the obligation to comply because it is the right thing to do or because they feel they have no other choice (Lee and McGovern, 2013, p. 107). The question of police power thus shifts from the police and the public belief in its legitimacy to the extent to which their actions can be justified in accordance with pre-established rules that govern society – these can be formal laws but also community values and expectations. Beetham's contribution to the legitimacy discussion suggests that the abstract and global theories about datafication and surveillance, which articulate a shift in what and who is seen as an authority in contemporary society, becomes relevant in the understanding of the productive nature of police power.

Debates on police legitimacy have been dominated by research that tries to understand why people comply with police actions, also known as audience legitimacy (Sunshine and Tyler, 2003; Tankebe, 2014; Tyler, 2006, 2003). Bottoms and Tankebe (2012) draw on key social and political science theories on legitimacy, most prominently Weber (1968) and Raz (2009), and their critiques to develop the 'dialogic model of legitimacy'. This model marks a theoretical shift away from a focus on audience legitimacy to a relational understanding of police legitimacy, a proposition made by the power holder, that is responded to by the public, which in some cases requires a response from the power holder. The model proposes two interrelated and interactive dimensions of legitimacy, that of power holder and audience legitimacy, which are in ongoing dialogue with each other. Similarly to media and surveillance studies, the authors foreground that there is a lack of empirical evidence on the operations of power-holder legitimacy – the police's self-belief, practices, and discourse – and argue that more research efforts should be directed at these operations to gain insights into how power holders build and maintain their claim to power. To understand power-holder legitimacy in the context of data-driven policing, I draw on Martin and Bradford (2021), who argue that the justification of power happens on two planes, that of micro- and meso-level legitimacy. Where the micro level refers to the individual police officer legitimacy claim on the streets and the meso level manifests in discourses, symbols, and norms around crime, justice, and professionalism. Meso-level legitimacy is seen to actively shape the image of the institution, also described as the 'corporate police voice' (Barker and Mulligan, 2003; Loader and Mulcahy, 2001). These insights point to a relationship between the productive nature of power and power-holder legitimacy, in which symbols and narratives are actively wielded by the police to justify their existence and actions.

Within the field of criminology, the justification of police power has primarily focused on research into audience legitimacy. Procedural justice has been a dominant approach to understand how the police claim to power is seen to foster normative obligations for the public to obey and comply with criminal justice actors and institutions. The origins of procedural justice in the context of policing can be traced back to psychologist Tyler (2006, 2007), whose interest lay in understanding what motivates individuals and communities to accept decisions in the criminal justice system. His starting point was in the proposition that legitimacy is a property that is possessed by the power holder that leads people to believe that the power holder's actions are justifiable and motivates people to obey their decision. There is a distinction between instrumental and normative modes of obedience: the first suggests that people comply with the law out of fear of punishment, and the latter relates to the understanding that the public complies because the police are seen as a legitimate institution whose actions are generally fair (Tyler, 2006; Tyler and Jackson, 2013). Tyler's disregard for instrumental modes stems from empirical research that showed how the threat of punishment in many cases did not incentivise people to obey the law (Kressel, 1985; McEwen and Maiman, 1981). A procedural justice effect manifests in the normative modes of obedience; the citizen assessment of the fairness of the procedures can explain why people voluntarily accept decisions they do not agree with and at times even perceive to be unfair.

Bottoms and Tankebe (2017) challenge the prominence of procedural justice in understanding why people comply with power. They argue that a more nuanced understanding of audience legitimacy is needed, one that accounts for the multiple publics the police need to address. 'That a single power holder may need to address several different audiences (the rich and the poor; different ethnic groups; etc.); also, audiences may make a significant differentiation in their assessments of the legitimacy of different power-holders' (Bottoms and Tankebe, 2017, p. 73). Here, they make two key observations: the justification of police power is contingent upon the way police communicate to distinct audiences, and the public's perception is influenced by their ethnic and socio-economic status and their exposure to and experience with the police. The procedural justice concept is too limited in scope, as it does not account for different perceptions of what just policing looks like. The authors do not offer an alternative to procedural justice but rather suggest that further legitimacy research 'needs to be connected more firmly to the literature on justice' (Bottoms and Tankebe, 2012, p. 158). Therefore, I will conclude my literature review by exploring different social justice theories to find an analytical lens that will allow this research to move beyond procedural justice in our understanding of what just of unjust policing looks like in a datafied society.

2.5 Data and social justice

In this final part of my literature review, I will engage with different theories of justice to find a framework that informs my research into how publics perceive the use of data-driven policing and its relation to power. Here, I will move away from the psychologist understanding of why people comply with police actions to a more sociological approach to justice. This is needed as procedural justice primarily tries to grapple with whether people find a specific policing action fair, and sociological approaches interrogate what fair procedures look like or should look like or, more fundamentally, what justice is or what it should look like in the context of policing. As such, I will now outline different sociological approaches to social justice to conclude that, rather than offering one ideal notion of just policing, I will follow Young's (2011) argument that justice is political and we should start by listening to 'new social movement', or what I call civic actors, to understand the injustices that materialise from the emergence of data-driven policing.

I will start by outlining Rawl's theory (1999) of distributive justice, which remains a dominant framework to engage with economic inequalities in society, and its critiques that open up space for other theories of justice to emerge. Rawl's (1999) argument starts from the premise that societies function by a set of rules that are recognised by people in their interactions with each other, while, for the most part, people act by these rules, these relationships are marked by conflict and personal interest. A set of principles are needed 'to provide a way of assigning rights and duties in the basic institutions of society and define the appropriate distribution of benefits and the burden of social cooperation' (Rawls, 1999, p. 4). Here, he proposes two principles of justice: the first assigns equal rights and duties to all citizens, and the second focuses on social and economic inequalities, where redistribution should ensure more equal distribution of primary goods in society. Resource-poor and marginalised communities should, according to Rawl, benefit the most from these acts of redistribution. The distributive justice model is based on the idea of a 'well-ordered society' where principles of justice are globally accepted and the mechanisms through which one can make a justice claim are considered to be adequate and acceptable (Cinnamon, 2017; Hoffmann, 2017).

Rawl's acknowledges that

> existing societies are seldom well-ordered and what is considered to be just and unjust is disputed. […] Those who hold different conceptions of justice can, then, still agree that institutions are just when no arbitrary distinction is made between persons in assigning basic rights and duties. (Rawls, 1999, p. 29)

He argues that, while there might be disagreements on what is considered just, if the institutions that govern justice claims make fair decisions, there will be no disagreement. This speaks to the notion of institutions as higher authorities that hold knowledge others don't to make decisions that are fair and just – an assumption that has been challenged by several scholars. Take, for example, Fontana (2008), who turns to Gramsci's notion of cultural hegemony to situate justice within the social norms and values that govern society. Cultural hegemony allows us to understand that, through the manipulation of cultural controls, the world views of the political and economic elite becomes the accepted cultural norm in society. Institutions as such cannot be seen separate from but as a part of this ideology that justifies the social, political, and economic choices that entrench political and capitalist power. Fontana (2008) points to the idea that distributive justice in itself does not engage with how justice is intertwined with capitalism and the contemporary political paradigm that create inequalities, it 'merely' aims to lessen the burden on those most disadvantaged within existing social and political structures. It is particularly this tension between Rawl's 'justice as fairness' and the injustice and exclusion resulting from the dominant political philosophy that has spurred critique from the feminist, anti-discrimination, sexual equality, and alternative globalisation movements (Young, 2011).

In her critique, Young (2011) argues that Rawl's theory of distributive justice does not account for the need for other forms of justice, like representation, recognition, restorative, transformational, and procedural justice, as it tends to dismiss non-economic wrongdoings, which are experienced by racialised, gendered, and impoverished communities. To understand these competing conceptualisations of justice, I turn to Fraser's (2008) theory of abnormal justice. 'Fraser advances a theory of justice that shifts our attention away from the dominant discussion on how goods should be distributed in a just society, and instead towards the very conditions that underpin how justice is understood, debated and advanced' (Dencik et al., 2018b). Fraser argues that public debates around justice assume a shared understanding of how justice is conceptualised, while in fact, the essence of social justice is up for grabs, as the 'what', 'who', and 'how' of justice have become contested in itself (2010, 2008). The 'what' of justice is disputed, as there is little shared understanding of the ontological claims underpinning the notion of justice. For the 'who' of justice, there is a disagreement about the kinds of actors who are entitled to make a justice claim in a specific context and whether this claim is even valid on a local, national, regional, and global stage. For the 'how' of justice, in a globalising world, there are conflicts about which institutions or procedures have the legitimacy to assess a specific justice claim and have the obligation to redress them (Fraser, 2008). Nancy Fraser's 'abnormal justice' forces us to move away from the notion of one ideal theory of justice, which strives to set out fundamental principles that should apply to most societies, and

move towards the acknowledgement of the multifaceted nature of our understanding of justice that at times conflict and overlap.

I now move to Sen's (2005, 2009) and Young's (2011) articulation of justice, as this allows me to situate Rawl's theory of justice and procedural justice as a top-down construction of what a justice claim could and should look like. Sen's (2005, 2009) capabilities perspective of justice argues that most theories of justice 'have tended to concentrate almost exclusively on the ideal of "just institutions" at the expense of an assessment of justice rooted in the actual lives that people are able to lead' (Dencik et al., 2018b). His capabilities perspective focuses on what a person is able to do or be within a certain context, through which he offers an approach that ties the concept of justice not to just institutions or just procedures but to the lived experiences of resource-poor and marginalised communities. In turn, Young (2011) argues for a more explicit political understanding of social justice. Instead of creating another ideal theory of justice, which is either inapplicable due to its abstract nature or is based on norms and values that are derived from the context in which the theory was created, a top-down diagnostic of social life, we should start by listening to the lived experiences of those who are affected by structures of power. In her theory of justice, Young (2011) takes her cue from the 'new social movements' of the 1960s and 1970s, which used the term *oppression* to point out the injustices that resulted at the hand of everyday liberal democracies (Young, 2011, p. 41). She argues that we 'should begin with the concepts of oppression and domination as it allows us to bring out issues of decision-making, division of labour and culture that bear on social justice but are often ignored in more philosophical discussions on justice' (Young, 2011, p. 4). Here, perceived neutral political and social structures are questioned as arrangements that enable and prevent justice claims. Thus, listening to social groups as a method will allow this thesis to move away from the normative notion of what justice ought to look like, often predetermined through top-down structures, and instead openly engage with how different publics perceive and experience the emergence of data-driven policing.

Data justice, the latest turn in discussions on social justice, is conceptualised to account for the ways in which datafication intersects with broader social justice concerns. Here, I will explore the different conceptualisations of data justice to argue that, while there might not be one single understanding of what is at stake, all contribute to an understanding that the datafication of society is invoking new kind of struggles over justice. Taylor argues that

> just as an idea of justice is needed in order to establish the rule of law, an idea of data justice – fairness in the way people are made visible, represented and treated as a result

> of their production of digital data – is necessary to determine ethical paths through a datafying world. (Taylor, 2017, p. 1)

It is exactly this idea of data justice that is debated. Take, for example, Johnson's (2016, 2014) account of information justice related to making social asymmetries entrenched in administrative data explicit. He argues that administrative data reflects dominant social norms and existing inequalities; when used in decision-making processes this in turn recreates and perpetuates privilege and injustice. Informational justice advocates for making these asymmetries visible (Dencik et al., 2019; Taylor, 2017). Heeks and Renken (2018) build on this idea of information justice and suggest that data can be a tool to accomplish greater distributive justice for the poor. Data for good, such as data collection on the Sustainable Development Goals, can be used to make the invisible visible, which allows for better decision-making and public accountability. In both accounts, the affordance of data relates to its perceived ability to expose injustices and inequalities, and it can be an instrument to improve the quality of life for all. This assumed correlation between visibility, transparency, and accountability has been long romanticised and problematised within social sciences (McGee, 2010; Grimmelikhuijsen, 2012; Ananny and Crawford, 2018; Cath and Jansen, 2021) and runs the risks of abstracting away or overlooking existing power structures that shape governance mechanisms (Selbst et al., 2019). In contrast, Dencik, Hintz, and Cable (2016) propose data justice as a lens to engage with who formulates a justice claim and argue that the dominant privacy and anti-surveillance discourse prevents datafication and surveillance from becoming a broader social justice issue.

Since these initial conceptualisations of data justice, scholars have offered different accounts of what is at stake with the increased datafication of society. Peña Gangadharan and Niklas's (2019) notion of 'decentering of technology' speaks to the distinctive origins of injustice claims, where the more tech-savvy civil society focus on technology and social groups 'prioritize the specific experiences of marginalized groups and "see through" technology, acknowledging its connection to larger systems of institutionalized oppression' (Peña Gangadharan and Niklas, 2019, p. 883). Decentring technology in an analysis of data justice should therefore offer much-needed nuance to the discussion about the role and place of technology in the production of social inequalities. Costanza-Chock's (2020, 2018) design justice speaks to how data justice can be achieved, where participatory design is seen as an empowerment tool through which communities get a voice in the development of data infrastructures that impact their lives. While these different accounts foreground that there is not one single understanding of data justice, it allows us to situate data in larger questions of power, oppression, and injustice. Thus data-driven policing should also be

approached as a site of struggle as it invokes new struggles over what injustice looks like and how the legitimacy claim of police is experienced and contested.

In these last two sections on the dialogic nature of power, I explored different social science frameworks that theorise about its productive and relational nature, where the police make a proposition of power that is responded to by different publics, which in some cases requires a response from the power holder. These debates inform my thesis by foregrounding that the police are in continuous negotiation with different publics to justify their claim to power. Any research into the relationship between datafication and police power, as such, needs to account for both power-holder and audience legitimacy and the interaction between them. This informs my research to approach data-driven policing as a study of practices, which should give me insights into police perception of power, and to study data as a site of struggle, where listening to social groups will inform my understanding of what just or unjust policing looks like.

2.6 Conclusion

A central concern of this thesis is understanding the relationship between data and power – more specifically, how the datafication of society introduces important questions about the nature of data-driven policing and its relationship with police power. In this thesis, I understand the concept of power as the ability of an actor to manage or control the actions of others despite their resistance, where the ability to attain and maintain power is dependent on the belief in an actor's legitimate authority and its ability to justify their actions. In this chapter, I position my research across multiple scholarly debates, predominantly within the fields of media and surveillance studies, criminology, and social justice. I discuss how these debates have informed my research and show the distinct areas that are still under-theorised. I build on research into data-driven policing to argue that data systems are increasingly becoming a medium through which the police engage with the world, yet there is still a lot unknown about its actual practice. I then turn to debates within the field of media and surveillance studies to understand how power dynamics are changing with the datafication of society. Here, media scholars offer ample insights into the ideological grounds of data and surveillance studies about the opportunity of contemporary data infrastructures to monitor people, objects, and events across time and space and use predictive analytics to control people's present and future behaviour.

These dominant debates within media and surveillance studies leave two things under-theorised. First, engaging with surveillance systems as a global phenomenon makes it hard to understand how

data-driven policing becomes embedded within an operational organisation like the police and how its introduction impacts local struggles for justice. Taking a cue from the empirical research on predictive policing and facial recognition, discussed in the first section of this research, contextualising global theories is imperative as it allows me to engage with how these data systems are shaping the discourse on crime, policing needs, and opportunities for action. Second, these debates primarily focus on how data systems change what and who are seen as an authoritative voice in contemporary society and how this allows different actors to manage people for political and market purposes. Building on debates in the social sciences, specifically in criminology and social justice, I argue that media and surveillance studies under-theorise about the productive and relational nature of police power. To conclude, in exploring these interconnected but distinct disciplines, it becomes clear that studying data as a practice and data as a site of struggle will allow me contextualise global theories of data and surveillance and theorise about the relationship between datafication and police power.

3. Methodology

In this chapter, I will elaborate on the theoretical foundations and methodological choices that inform my research. In the previous chapter, I observed how there is still a lot unknown about the turn to data-driven policing in Europe. To answer my research questions – what is the nature of data-driven policing? What is the relationship between datafication and police power? – I build on scholarly debates that approach the emergence of data-driven policing as a social process. This entry point into police's turn to data allows me to research data as practice, what Couldry (2004) has called studying 'media as practice', to engage with data-driven policing as a sociotechnical system that is shaped by both the technology and its surroundings. This entry point requires an articulation of how I approach the relationship between society and our knowledge of it. This thesis roots itself in the broader tradition of critical social sciences, more specifically critical realism, which views knowledge about society as socially constructed, contextual, and stratified. While true knowledge of social structures is beyond the direct experience of people, this philosophical orientation argues that analysing the use, discourses, experiences, and perceptions of data-driven policing will offer insights into the relationship between datafication and police power.

In this chapter, I will outline the theoretical foundations and methodological choices that inform my research. First, this chapter will engage with critical realism as a philosophy of study, specifically its structure and agency debate, which offers the theoretical grounding to study how power and counter-power are enacted and justified within social structures. Second, I will elaborate on the choice for an empirical inquiry into the praxis of police and civic actors in relation to data-driven policing as an object of study, offering the justification for my methodological choices on how knowledge is produced. Finally, I will engage with the operationalisation of researching praxis, my data collection methods, samples, and the limitations.

3.1 Critical realism

In its concern with data-driven policing, this research is interested in understanding the relationship between datafication and police power. Here, police power is approached as a continuous negotiation of what fair and just policing should look like, and research is conducted with the aim to 'expose and explain power structures and relationships with the view to alleviate unnecessary and unwanted suffering' (Fay, 1987). As such, this research is rooted in the broader tradition of critical social science, and more specifically in critical realism, which approaches knowledge about society as socially constructed, stratified, and contextual. In this first section of my methodology chapter, I

will engage with critical realism as the theoretical foundation that informs my research, and its structure and agency debate positions agency both as an act in itself and a response to the social structures that shape society.

Critical realism, closely associated with Roy Bhaskar, was developed as an alternative to positivism and conventionalism (Bhaskar, 2008; Dean et al., 2005). Bhasker argues in his book *A Realist Theory of Science* (2008) that science should start from what is knowledge (ontology) instead of from how is knowledge possible (epistemology). He argues that a focus on the 'how' might run the risk of observing the visible and overlooking the less visible structures and mechanisms of society, which will restrict our understanding of reality (Danermark et al., 2002, pp. 5–6). Bhaskar's critique on positivism is twofold: first, he argues that it is reductionist, reducing reality to something that can be directly observed. Jessop (2005) calls this Bhaskar's anti-positivist naturalism. 'Baskhar argues that, while hermeneutics claims correctly that the social world comprises a pre-interpreted reality, it does not follow that the social world is reducible to the ideas that people have about it' (Jessop, 2005, pp. 41–42). Second, Bhaskar critiques the positivist presumption that methodologies used in natural science, where the object of study is naturally produced, passive, and unchangeable, can be applied to social sciences, in which the object of study is ever-changing and messy.

This critique of positivism allowed Bhaskar to develop critical realism as a philosophy that adopts ontological realism and epistemological relativism. Ontological realism insists upon a world, and its structures and power relations, that exists independent of human knowledge or human perception of it. Knowledge about society is stratified; the object of knowledge is the real, which is beyond direct experience and can manifest itself in events and mechanisms, the actual, which are experienced on the empirical level (Danermark et al., 2019, 2002; Easton, 2010; Jessop, 2005; Sayer, 2010, 1999). It is the objective of critical realism to 'fill[s] the gap between the real and the experiential by attempting to attain true knowledge of specific powers or mechanisms as these are located in the different layers of which nature is composed' (Dean et al., 2005, p. 8). This stratified view of society allows me to engage with data-driven policing projects as a manifestation of social structures and mechanisms that create conditions for its emergence. Thus, rooting my research in critical realism allows me to study the emergence of data-driven policing to understand what is changing in society, in this case how data is transforming the exercise and justification of power and counter-power.

In the critical realist position, it is important to recognise that, unlike the natural world, the social world has displayed few constants; the fabric of society and its structures and mechanisms have

changed over time and space. Here, it is argued that what is true in one context at one moment in time does not necessarily have to be true in another context at another moment. The view that knowledge of society is temporally and spatially determined, as such, provides the foundational argument for the need for contextual research (Danermark et al., 2002; Jessop, 2005). Critical realists argue that, when research finds a strong correlation between events, it is important to specify under which conditions this might be the case, as other forces might have influenced the relationship. For example, if crime reduction is witnessed after the introduction of a predictive policing programme, this might indicate a direct correlation; however, a critical realist reflection requires us to take other spatial and temporal elements into account, such as dropped prices for electronics, improved socio-economic conditions, or broader investments in social welfare services (O'Mahoney and Vincent, 2014). The understanding that the process of scientific knowledge production is contextual and can be overturned in the (near) future is what critical realists refer to as epistemological relativism (Dean et al., 2005, p. 8). This does not imply that all knowledge is up for grabs or that there is no truth, but rather that some research and knowledge will be more accurate than others, and through iterative research practices, a critical realist will slowly come closer to this possible truth (Dean et al., 2005; Zachariadis et al., 2013). This informs my approach of contextual and iterative research to account for both the data systems and social-political dimensions that influence how practitioners and civic actors come to understand crime, police power, and justice.

I will now outline critical realism's prominent structure and agency debate as it informs my understanding of agency in the context of police power and datafication. Within critical realist discussions, structures pre-exist and establish the conditions for people's lives, in which they create and restrict conditions for actions. Structures consist of socially created concepts such as norms, cultures, economic systems, race, gender, and class (Danermark et al., 2019, p. 11). Agency can be found when people exercise their power to act, to reproduce, or to transform the condition of their lives. Critical realists argue that we need to make an analytical difference between structure and agency, to approach them as two separate phenomena that are mutually dependent, and to study the interaction between them over time (Danermark et al., 2019, 2002). This informs my research approach, in which I study the praxis of data-driven policing and that of civic actors to shed light on how data systems are perceived to transform police power in European societies. Within the structure and agency debate, there are different schools of thought as to what enables and constrains agency. There is a general agreement that research into power relations should take the positionality of actors into account, but scholars offer different interpretations on the extent to which these actors can act outside predetermined rules of action.

Giddens's (1984) structuration theory engages with the concept of agency, not in relation to the act itself but to the ability of an individual to act in the future, where people have the power to act and act differently in each situation. Yet the dominant normative construction of how one ought to behave in society sets the rules for action (Danermark et al., 2019, p. 77; Giddens, 1984). Archer's morphogenetic approach (2013, 1995) and Jessop's strategic-relational approach (2005, 2001) move beyond structuration theory and argue that agency as causal powers are both shaped by social structures and by agents' ability for reflexive deliberation. Archer's (1995) model argues that structures favour certain actions over others, and the specific position agents occupy within these structures, as well as their resources and capacity to reflect on their actions, will influence their ability to act. The closer actions align with the normative expectations of how one ought to behave in society, the less friction and conflict there will be. Jessop builds on this observation and argues that structures should be treated as strategically selective, where specific structures and specific structural configurations 'privilege some actors, some identities, some strategies, some spatial and temporal horizons, some actions over others' (Jessop, 2005, p. 48). Actions are treated as structurally constrained, meaning that actors have the ability to reflect on their position in a structure and the success and failure of their action and strategies thus far, and have the potential to make strategic choices to adapt their actions. In this chapter, I build on Elder-Vass's (2010) argument, which tries to reconcile these distinct approaches by arguing that agency is both determined by the social structures and reflexive deliberation, as an individual can be highly reflective in one area but uncritically reproduce attitudes and behaviours in other areas.

Finally, I want to draw attention to Jessop's (2016, 2012) rejection of the state and state powers as neutral instruments or benevolent agents. He argues that we need to understand state power and its institutions within its asymmetries of authority and domination, 'with its structural and strategic role in reproducing wider patterns of exploitation, oppression, and domination at particular times and in particular places; and with the scope for challenging, modifying or overturning these asymmetries and their effects' (Jessop, 2016, p. 3). This observation requires us to engage with the police and its practitioners as agents with the power to reproduce and transform structures of oppression and domination but who also have the ability to challenge and transform them. Jessop (2016) further argues that state institutions do not operate in a vacuum; they are also subject to dominant social norms and the influence of individual politicians, the government, and parliament, actors who set priorities, budgets, and expectations for and on them. Any research into the police thus needs to account for the notion that a complex society has several state institutions that collaborate and compete with as well as reinforce each other. Thus, to examine how data systems are perceived to transform police power, the critical realist approach demands contextual research

that explores the praxis of both police and civic actors in relation to data-driven policing and that of other stakeholders.

3.2 Research design

> Social reality swarms, it flows, and rushes, it is rich in levels and complexity, consisting of many powers and influences, events, and experiences. How is it possible to research such a mess? (Danermark et al., 2019, p. 71)

This quote draws our attention to the challenge of producing knowledge about the messy nature of reality. In line with the critical realist approach, which observes that the object of study informs the type of knowledge that can be produced and the methods that are most suitable for it (Danermark et al., 2019; Easton, 2010), I will start this section on my research design by first defining my research field and empirical research sites. Then I will outline my data collection methods and sample, and I will conclude with the methodological limitations of my research.

In my literature review, I observed how dominant scholarly debates in the field of media and surveillance studies engage with data systems as a global phenomenon that allows power holders to control the present and future behaviour of people for political and economic purposes. My topical analysis showed a number of knowledge gaps; there is still a lot unknown about the actual nature and practice of data-driven policing (Brayne and Christin, 2021), and its introduction is invoking new struggles over justice (Dencik et al., 2019; Peña Gangadharan and Niklas, 2019). In addition, in global surveillance theories, the productive, relational, and normative dimension of power that becomes enacted and embodied through data systems remains under-theorised. The scientific debates discussed in chapter 2 informed the delineation of my research field, where an inquiry into data-driven policing is a study of social processes. Here, I build on Couldry's (2004) studying 'media as practice', to argue that exploring the social dimensions of the police use of data requires a study of practices. This allows me to engage with data-driven policing as a sociotechnical system that is shaped by both the technology and its surroundings and to account for the imaginaries, discourse, conflicts, and struggles that occur when technology becomes embedded within the police. I choose to engage with the study of practices through qualitative research methods and structure my empirical work around case studies to produce a few detailed descriptions of concrete events to identify underlying causal structures and generative mechanisms (Ackroyd and Karlsson, 2014; Easton, 2010; McKechnie, 2007).

My empirical research consists of a mapping, two cases studies, and a civic actor chapter in the context of Belgium, Brussels, the Netherlands, and the UK. I will briefly introduce each of the empirical sites of research before outlining my data collection and sampling methods. As observed in chapter 2, there is still a lot unknown about the turn to data-driven policing in the context of Europe. Therefore, I start my research with a mapping chapter to understand what is actually happening, by exploring where police are using which types of data systems. This will allow me to develop a bird's-eye view of what police are investing in, gain a sense of external dynamics that shape the conditions for data-driven policing to emerge, and identify events that are representative of the changing nature of the real, which provides the basis of my case study selection. My mapping shows how most projects are aspirational and either have not materialised in the day-to-day of policing or have since been halted; as such, I found that the use of specific data-driven policing is a nascent practice that is ephemeral in nature. Furthermore, I found that there are only a small number of practitioners working in this niche field, who are either in senior management functions, where their job revolves around introducing data and technology within the organisation, or are specialists that develop a specific tool in relative isolation. This has consequences for how we can study data as practice and informed my case study selection to focus on specific functions that are developed simultaneously across geographies over that of focusing on one specific implementation. Engaging in multi-sited empirical research, in which I looked across and between data-driven policing functions in different jurisdictions, allows me to identify broad organisational principles that structure police approaches to technology and offer insights into the social structures that (re)produce them. I use the term *function* to describe specific data affordances (Gibson, 1979; Hutchby, 2001) that are ascribed to a suite of data-driven policing technologies.

For my case studies, I selected the function of data-driven risk scoring and biometric recognition. I use the term *data-driven risk scoring* to refer to the practices of using police data to infer risk scores on the likelihood that an individual will become a victim or perpetrator of a specific criminal offence. The term *biometric recognition* refers to the processing of the digital representation of face and voice features for the purpose of a criminal investigation. These functions emerged from my mapping study as sites of interest where the police actively invest in and develop tools with the aim of introducing them into the core of policing to ensure safety and security on the street. Researching data as practice allows me to explore why police are turning to these functions; how it becomes embedded within an operational organisation like the police; how practitioners frame its use, policing needs, risk, and opportunities; the challenges they experiences; and the practices that emerge from the introduction. As observed in my literature review, this is only part of the story; police power is a continuous negotiation between the enforcement of specific rules through the

threat of punishment and civic responses to it. Therefore, studying police practice will only offer a partial view on how the relationship between datafication and police power comes into being. My final empirical site is that of civic responses to data-driven policing, in which I listen to a range of civic actors to understand how data-driven policing is shaping debates on what just and unjust policing look like.

3.2.1 Methods and data

Critical realism does not prescribe a specific research method and allows for appropriate data collection through a mixed-method approach. Research on developments in policing is notoriously difficult in terms of access (Brayne and Christin, 2021), even more so when looking into the use of data systems, as only a small number of senior practitioners are working on it and some are hesitant to speak to outsiders on projects they feel have become politicised in recent years. Taking these challenges into account, I explore the practice of data-driven policing by drawing on a mixed-method approach to provide insight into the uses and perceptions of data-driven risk scoring and biometric recognition. In this section, I will describe my data collection methods, primarily semi-structured interviews that are substantiated with participant observation and the study of grey literature, and my sampling approach for each of my three empirical research sites through which I answer my research questions.

3.2.2 Interviews

> Interviews have been described as both an art and a science. (Robinson, 2013)

My primary method of data collection is semi-structured interviews, which allows me to listen to the experiences, discourses, and beliefs of experts, police practitioners, and civic actors to create knowledge on people understanding and interpretation of data-driven policing and its relation to police power. I draw on Robinson (2013) and Smith and Elger (2014), who argue that this method is particularly well suited to study power relations in society, as interviewing allows people to explore and reflect upon their practices and the broader environment within which they are situated. Traditionally, a distinction is made between three types of interviews – structured, semi-structured, and unstructured interviews. In the first, researchers work with a list of predetermined questions, which are asked to all participant in the same order. Semi-structured interviews involve asking a series of predetermined, open-ended questions, with space for follow-up questions. And unstructured interviews offer a loose framework in which the research allows the interview to go wherever the participant wants to take it (Brennen, 2014; McKechnie, 2007). This research has

chosen to apply semi-structured interviews to collect data from different perspectives on the same topics and still have the freedom to dive deeper into specific observations and experiences. Below, I will explore my sampling approach for interviewing across the three sites of empirical research: mapping, cases, studies and civic responses.

In my research design, I'm cognisant that sampling is a question of power in itself; it determines who is privileged to speak, to tell a story, and to be heard in research (Jeppesen and Sartoretto, 2020). To account for these power dynamics, I engaged in an iterative research and sampling approach and made deliberate choices to include a wide range of voices within my research. I start with a mapping of what is actually happening in Europe, which informed the selection of topics and locations for my case studies, and in turn narrowed the civic actors to distinct geographies. As such, my interview processes is characterised by three phases: expert, practitioner, and civic actor interviews. The expert interviews functioned both as interviews, reflecting their reality outside the interview, and as insider references to gain access to police practitioners for my case studies (McKechnie, 2007; Patton, 2002), and the practitioner and civic actor interviews primarily functioned simply as interviews. Here, the samples for my three phases of semi-structured interviews are a combination of purposive and snowball sampling (Etikan, 2016; Patton, 2002; Tongco, 2007).

Purposive sampling, drawing on insights from grey literature, I identified and reached out to a number of key experts for my mapping study. These interviews provide the basis of my mapping chapter, but, where relevant, this mapping chapter also includes observations from other interviews. The findings of my mapping informed the second phase of my research, the case study sample, where I searched for people involved in the identified data-driven risk scoring and biometric recognition projects. I reach out to those identified practitioners with a request for interviews, where possible through existing networks or otherwise by sending a cold email. My civic actors' sample was informed by Browne's (2015) observation that race is a central organising principle for surveillance, and as such, I tried to include a diverse range of voices in it, from those that are working on the issue related to data-driven policing in digital rights and human rights organisations to grassroots actors who are working on police violence and the racialisation of crime. Snowball sampling means that, after establishing contact within the different police forces, civic actors networks, oversight bodies, and other government agencies, I asked for insider references to gain access to the people that are most difficult to gain access to, in other words, police practitioners and civic actors working on the front line of anti-discrimination work. Furthermore, at the end of my interviews, I asked each person if they could suggest and connect me to other people to interview.

Throughout my research, I conducted fifty-six semi-structured interviews,[2] with twelve experts, twenty-four police practitioners, and twenty civic actors (see figure 1). In this section, I will outline my sample for each of my three empirical research sites. For my mapping, I conducted interviews with police experts from Belgium, Brussels, the Netherlands, and the UK. I define police experts as police officers, regulators, members of police oversight bodies, and technologists who are actively working in the area of data and police. The sample was chosen based on interviewees' different expertise, their bird's-eye view on what is happening in the national or European context, and the geographical spread. I narrowed down on these three countries from the observation that the police is actively experimenting with a range of data-driven policing functions and they have similar 'just-do-it' policing cultures when it comes to the introduction of new technologies. Brussels is included in my sample as the LED (EU Directive 2016/680[3]) became enforced at the start of my data collection process, which set the data protection standards for European police forces. In addition, its funding programmes, Horizon 2020, and its predecessors have actively invested in policing technology development under its security research projects.

For my case studies, I interviewed police practitioners, police managers, city officials, and relevant oversight bodies who are actively working on data-driven risk scoring and biometric recognition. My sample consisted of senior police officers, who are either in charge of the district in which the data-driven policing function was being developed, lead the biometric units in the forensic department, are responsible for innovation within the police, or are the specialists that develop a specific tool in relative isolation. The sample for my case study on data-driven risk scoring is informed by a number of insights. My mapping study showed that primarily the Dutch and British police are actively developing and deploying a number of projects; there, many of the practitioners who work on risk-scoring projects, which received a lot of negative press in recent years, did not respond or responded negatively to my request for interviews. The choice to include city officials in this practitioners sample stems from the observation that data-driven risk-scoring approaches are often implemented in cooperation with a range of other state institutions and are at times even housed within the municipality. For my biometric recognition case study, my mappings informed my choice to focus on the developments in Brussels, Belgium, and the Netherlands. At the time of my interviews, the use of facial recognition by the police in the UK was under increased public

[2]One interview was off the record: the participant was comfortable with me writing down notes, but I was not allowed to record or directly attribute findings to them as an identifiable person.

[3]Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection, or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L 119/89 ('LED').

scrutiny, which affected my ability to speak to police practice on the record. British practitioners I reached out to, through warm and cold contacts, either declined my request for interviews or refused to speak on the record. I did include an interview with a Danish police practitioner, as he is vocal within the European policy discussion on police, technology, and privacy.

Semi-structured interviews			
Phase	**Country**	**Type/Title**	**Location**
Twelve expert interviews	Brussels, Belgium, the Netherlands, and the UK	Police lead, police managers, police adviser, regulator, police oversight, and technologists	Belgium federal police, Dutch national police, UK police, public prosecutor, European Commission, oversight bodies, and independent technologist
Twenty-four police practitioner case studies	Brussels, Belgium, the Netherlands, and the UK	District chief, police managers, chief superintendent, detective superintendent, chief information officer, lead police data scientists, data protection officer, policy adviser, chief privacy officer, lead biometrics specialist, policy adviser commissioner general, biometric director, and relevant oversight actor	Amsterdam municipality, Amsterdam police, Gelderland-Midden police, Noord-Holland police, Dutch national police, West Midlands Police, Hampshire police, Belgium federal police, Interpol, Danish federal police, Police and Crime Commissioner's office, Controleorgaan op de politionele informatie, Authoriteit persoonsbescherming, and the Information Commissioner's Office (ICO)
Twenty civil actors	Brussels, Belgium, the Netherlands, and the UK	Executive directors, European policy director, policy adviser, senior advocacy officers, racial justice advocate, community organiser, and legal advisers	a.o. Access Now, Ada Lovelace Institute, Amnesty International, Big Brother Watch, Bits of Freedom, Digital Freedom Fund, Edri, ENAR, Liga voor de Mensen Rechten, PICUM, PILP NJCM, State Watch

Figure 1: Overview of semi-structured interviews

The third phase consists of interviewing civic actors, defined as both formal civil society organisations working on issues of digital rights, human rights, anti-discrimination, ethnic profiling, and police brutality, and also individuals whose repertoires of action to change the social structures that determine life chances of those targeted by data-driven policing practices in Belgium, Brussels,

the Netherlands, and the United Kingdom. This sample was chosen to gain an understanding of the injustices, experiences, strategies, and responses to data-driven policing (Coole, 2005; Danermark et al., 2019, p. 93). The geographies were chosen to match the empirical research site of the mapping and the two case studies.

For each of the three phases, I created a list of open-ended questions (see appendix I for example interview questions). The questions are informed by my topical literature review, my research questions, and my methodology, and they inquire into specific topics, such as the use of a data-driven policing function, the practices, oversights, and safeguards, and concerns and injustices (Brennen, 2014; McKechnie, 2007). As a general introduction, I asked all interviewees to describe their job, position in the organisation, and how they are involved in data-driven policing. In my mapping research, I asked about the trends in data-driven policing, if they knew of any specific tools that are being used, where, and who was involved. For my practitioner interviews, my questions related to the actual tool, its origin story, the development process, and its current status. In my civic actor interviews, I asked participants about which data-driven policing projects they were aware of, what their concerns are with these, and asked them to respond to some practitioner claims, for example, ‘in the end, a judge will decide if it was fair or not’. I asked all of them about the needs, opportunities, challenges, and risks related to data-driven policing, as well as about their perception of oversight and safeguards. In my practitioner and civic actor interviews, I also asked participants to reflect on their allies or biggest critiques.

Prior to each interview, I researched the project, participant, and their organisation, which informed the questions I prioritised in the interview. The interviews were conducted in Dutch and English, lasted between sixty and ninety minutes on average and, when possible, were conducted face to face at the participant’s place of work. Due to the Covid-19 pandemic, half the interviews in my biometric recognition case study and all civic actors interviews were conducted online through Zoom, with both the audio and video turned on. Before each interview, I shared the scope, intention, and goal of the research and obtained consent for the recording, analyses, and quotations (see appendix II for information sheet for potential interviewees). Several practitioners and civic actors indicated that they did not want to be quoted by name and some asked to ensure that they could not be identified. Considering there are only a small number of practitioners and civic actors working on data-driven policing, it is relatively easy to re-identify research participants; as such, I have chosen attribute quotes using basic job descriptions, and only where it would add value associate the geographical location to them. Each interview was digitally recorded and transcribed verbatim, and quotes were cleaned and, where needed, translated into English at the writing stage. For the

cleaning and translation, I tried to stay as close as possible to the actual words used by practitioners. However, the Dutch practitioners and civic actors used a lot of proverbs, sayings, and slang to refer to certain developments; in these cases, I privileged the meaning of what was said over a literal translation.

I chose to manually code my interviews and thematically analyse them to best capture the experiences, meanings, and ideologies of my interviewees. I systematically searched, identified, and analysed interview data for common themes to offer insights into my research questions (Boyatzis, 1998; Braun and Clarke, 2006; Brennen, 2014). In my coding process, I looked for specific projects, tools, and developments; their origin story and uses; the actors involved in the projects; the discourse on their needs, risks, challenges, questions, and points of conflict; and the practices that emerged from the introduction of data-driven policing. Critical realism informs my thematic analysis on two accounts. First, my role as a researcher, where any analysis is not a passive account of reality but rather an interpretation of the interpretations of others, what critical realists call 'double hermeneutics'. This requires researchers to be aware of their own biases and systematically and transparently code the interviews. Second, critical realism argues that insights on the mechanisms and structures that shape society are obtained through a process of continuous and iterative knowledge creation. As such, the coding of interview data is conducted through an iterative process in which my starting point is to let the data speak for itself, identifying the themes from the interview data rather than having the themes informed by theory. The choice to engage in a data-driven analysis informed my research question to evolve from a broad question – what is data justice in the context of policing? – to a more specific ones – what is the nature of data-driven policing? And what is the relationship between datafication and police power?

My coding process is based on Braun and Clarke's (2006, p. 87) phases of thematic analysis, which involves a constant moving back and forth between interview data, codes, and analysis. After each phase of interviews and after an initial familiarisation of my data set, I started generating initial codes by highlighting specific data and leaving comments in the margins. This allowed me to organise my data and start to identify common themes (Boyatzis, 1998; Tuckett, 2005). After I coded all the interviews from all the interview phases, I went back and recoded them. This allowed me to include insights and themes that emerged during the coding process across my entire data set. In the third phase, I grouped all codes and underlying data into a potential theme (Braun and Clarke, 2006). In a separate document, I clustered, grouped, and regrouped codes into possible themes and further refined the clusters by rereading all of them, seeing if they form a coherent pattern, and weighing their prevalence in relation to my overall research questions. Subsequentially, I reread my

interview data to see if my themes reflected the meaning evident in the entire data set and identify whether there was any data missing in my earlier coding. I finished by refining the themes and underlying data, making selections of compelling quotes that best illustrate my findings, and relating the data to my research questions and literature.

3.2.3 Observations and grey literature

As previously noted, research on developments in policing is notoriously difficult in terms of access (Brayne and Christin, 2021), even more so when looking into the use of data systems; as such, I have chosen to substantiate my interview data with participant observation and the study of grey literature. These secondary research methods are used to enrich the findings from my interviews (Ackroyd and Karlsson, 2014; Danermark et al., 2019). I will first discuss the method and sample of participant observation and then that of studying grey literature.

Participant observation enrich my research as it offers additional insights into the discourse, attitudes, activities, and dynamics of and between police practitioners and civil actors (Aktinson and Hammersley, 1998; Brennen, 2014; Silverstone et al., 1991). After my initial expert interviews, I was invited to both sit in on police and civic actor meetings on data-driven policing. These invitations ranged from closed-door meetings organised by police to discuss a specific risk-scoring project and the use of biometric recognition to interdisciplinary meetings where police, civic actors, and oversight bodies came together to discuss the general phenomenon or a specific function. I also attended closed-door civic actor meetings on mobilising, campaigning, and litigation tactics to halt the turn to data-driven policing. My sample was determined by the relevance of the meeting to the topic and the geographies of my case studies or civic responses. For my participant observation, I chose to be an observer as participant, rather than a complete observer (observer from a distance), participant as observer (not fully integrated into the culture), or complete participant (going native) (Brennen, 2014). The role as observer as participant, in which the researcher is on-site but not fully integrated into the culture, allowed me to observe and take notes about what was discussed, how these issues were discussed, who was seen as the primary experts, and the dynamics and invisible hierarchies between those present at the meetings. Although I was not part of the conversation, this method left space for interaction, conversation, and follow-up interviews with individuals who were present at the events. I documented my participant observation in what I call field notes on my computer. During the meetings, I documented who said what, observations on themes that were responded to or ignored, items I wanted to follow up on, and anything that caught my eye in the moment. After the meeting, I wrote down a reflection of the meeting, how issues were discussed,

who was in the room, who was missing, and the dynamics between the actors. Appendix III offers an overview of the meetings and events in which the participant observations were conducted in the period of 2018–2021.

My fieldwork is further substantiated with the method of studying grey literature, in other words, reports, documentation, policy reviews, police presentations, and media reports. In academic debates, the term *grey literature* emerged to account for knowledge that is embedded within written text produced on all levels of government, academics, business and industry in electronic and print formats not controlled by commercial publishers (Auger, 1998). Grey literature primarily substantiate my mapping chapter through snowball sampling. I looked at forty-four police, oversight, government, and civic actors websites in Brussels, Belgium, the Netherlands, and the UK for information about the use of data-driven policing either by searching for specific technologies or functions on the website. These websites offered direct information about the programmes or were a springboard to thirty reports on aspects of data-driven policing. Where needed, I searched for media clippings on a topic that emerged from the different websites and reports to offer more insights into specific deployments. Figure 2 offers an overview of the types of grey literature that I engaged with for my mapping chapter. After reading the initial reports, documents, and blogs, I looked at key references and continued to read those.

Type	**Number**
Police websites	14
Government websites	6
Oversight website	8
Civil society websites	16
Police, government, and oversight reports	16
Civil society reports	14
Police and security investment portals	2
Legal documents	2
News outlets	20

Figure 2: Overview of grey literature mapping study

For my case studies, I engage with grey literature on the specific data-driven policing tools, which includes information on police websites, evaluations, training manuals, videos, and other materials that are created by and for the police. These documents give insight into the when, what, and how of a specific data-driven policing tool. If and when participants referred to a specific document,

Freedom of Information request, policy review, or report, I would try to gain access to it. Rothstain and Hopewell (2009) argue that, when engaging with grey literature, it is imperative to be aware of the subjectivities that are embedded within it and the sample bias of finding information that is most easily available. For my research, the subjectivities embedded within grey literature are part of its appeal; these are manifestations of the discourse, attitudes, priorities, and praxis of a specific group, and the availability is in part a reflection of which voices are shaping the discussions on data-driven policing.

Through the process of deliberate and snowball sampling in my interviews, in combination with participant observation and the study of grey literature, I reached theoretical data saturation (Fusch and Ness, 2015; Guest et al., 2006). The question of when I collected enough data was primarily informed by the fact that the mixed methods no longer provided fresh insights. In my case studies, theoretical data saturation was informed by the number of people developing data-driven policing functions who were willing to speak on the record and the observation that the last interviews no longer provided new insights that I had not already gained from my overall data collection process.

3.3 Methodological limitations

Bhaskar (2008) argues that any philosophy of science has to resolve itself with the fact that knowledge is a social product. Individuals produce knowledge and its process will always be influenced by people's own situations, experiences, and biases; as such, our perspectives are always situated someplace. This understanding of the subjectivity of knowledge production requires some reflection. First, centring data-driven policing in our understanding of police power privileges data as a transformational power over other social, economic, and political factors. This runs the risk of displacing other structures such as race, class, and gender in our understanding of contemporary police power. Second, my positionality as a researcher will to some extent influence what is considered relevant knowledge, as the study of social objects is one of double hermeneutics, in which I observe and interpret the experiences and interpretations of others to create an understanding of the world (Jansen, 2020). This does not imply I cannot find or come close to the truth; from a critical realist perspective, it requires me to account for the possibility that bias can occur in the selection of cases studies, methods, interview samples, events, and analysis (Zachariadis et al., 2013, p. 859). Therefore, this research is theoretically informed; it started with a mapping study that informed the selection of my case studies and sample, and I strived to collect data from a wide range of perspectives.

Knowledge and the finding of the truth is not only limited by the researcher's context, it is also influenced by the preinterpretation of others. For the latter, it is important to recognise that my research department has no institutional relationships with the police, and while the combination of purposive and snowball sampling allows me to include a wide range of police practitioners' views, these are also the ones interested in talking to an outside researcher. In my civic responses to data-driven policing, I chose to listen to civic actors for the formulation of injustices and not to the voices of targeted communities, a decision that was informed by the fact that most data-driven policing efforst are still in their infancy and many are not yet deployed. In addition, this research did not want to be extractive in nature by dropping in and out of communities for the mere purpose of knowledge production (Pacheco-Vega and Parizeau, 2018). However, I realise this choice inscribes power to civic actors over that of lived experiences and takes a position on how injustices are formulated (Young, 2011; Zuberi and Bonilla-Silva, 2008). In further studies, it is imperative to work with and alongside targeted communities over a longer span of time to define the problem and also pathways forwards.

The use of semi-structured interviews has the additional challenge that data is collected through the social interaction of the researcher and participants. Seale (1998) identifies two major traditions on which the analysis of interviews has centred: interview data as a resource and interview data as a topic. Where the latter understands interview data as a reality jointly constructed by the researchers and participant, and the former sees it as reflecting the reality outside the interview (McKechnie, 2007). In this research, it is important to be aware of both the influence the researcher can have during the interviews and when analysing them. To limit researcher bias, I presented the research, questions, and body language as neutrally and openly as possible, and I was conscious of a diverse and well-represented sample in which different voices can speak. To minimise possible bias in the interviews, the themes have been theoretically informed and have been selected in relation to the research interest. During the interviews when something was unclear and there was a need for further explanation, follow-up questions were asked on specific issues, experiences, and discourses. Facts that are mentioned in interviews were verified and triangulated with other interviews, grey literature, and participant observation.

The process of interviewing is characterised by a power relation between the researcher and the participant. Building trust and accounting for possible bias and microaggressions is particularly important when interviewing police practitioners and civic actors. Practitioners defined the police as a closed shop who, in general, are less open to speak to outsiders. In addition, there is no doubt that when interviewing actors who work on anti-discrimination issues and are part of or working closely

with targeted communities, who are predominantly persons of colour, my positionality as a White Dutch female researcher will manifest itself in the interview protocol and the logic of this research (Zuberi and Bonilla-Silva, 2008). Therefore, I tried to create a personal connection with the participant prior to each interview to create familiarity and a trust relationship.

Third, the interviews and analysis are conducted in accordance with Cardiff University's Research Integrity and Governance Code of Practice (Cardiff University, 2019) that, among others, stipulates how to ethically conduct research and store data. In addition, my practice has been informed by the Do No Harm principles (Engine Room, 2016), which stem from the belief that all actions and behaviours lead to consequences, both positive and negative. Through responsible behaviour and data processing, I seek to increase the positive impact of my actions and reduce possible negative impacts. The final limitation is in the generalisability of this thesis, which plays out at different levels. The critical realist position is that knowledge is created within a demarcated spatial-temporal horizon, which has consequences for the extent to which findings from one time and place can be generalised to another. Here, the value of case studies as a method lies in the ability to uncover complex power relations in society; however, it also impacts the ability to generalise findings to other structures and contexts. Interview as a method 'offers insight into human interaction, relationships, and meaning creation, and an examination of social processes and forces at work in culture, economic, politics and other aspects of society' (Robinson, 2013), but offers a smaller sample size than quantitative research practice. These observations do not mean that certain abstractions and generalisations are not possible. It merely means that some findings might be rejected in another context in the future, and others will contribute to the broader understanding of the possible truth.

3.4 Conclusion

This chapter outlined and justified the theoretical and methodological approach I build on to explore the actual nature of data-driven policing and the relationship between datafication and police power. I position my research within the philosophy of critical realism, which views knowledge about society as socially constructed, contextual, and stratified. More specifically, I draw on the prominent critical realist structure and agency debate, which informs my thesis's understanding that exploring data as practice will allow me to uncover how datafication is redrawing the lines on which power and counter-power becomes enacted in society. In my research design, I explain my choice for case studies and the use of mixed methodologies, in which the primary form of data collection consists of semi-structured interviews that are substantiated with participant observation

and grey literature. I make the argument for my three empirical sites of research – mapping, case studies, and civic responses – to add much-needed evidence to the debate on the relationship between data and police power. Throughout this chapter, I have explained my theoretical grounding, methodology, case study selection, sampling approach, and their relevance for answering my research question. Additionally, I have engaged with questions of power that manifest when data is centred as an object of study, in the sample of interview participants and in whose voices are foregrounded in this work. I concluded this chapter by engaging with the limitations of my research.

4. Mapping data-driven policing in Europe

In line with the critical realist tradition, my research starts from the notion that policing practices are shaped by and shape their social and political contexts, which requires an inquiry into data-driven policing to be contextual. I found that dominant scholarly debates that explain and critique the police turn to data-driven policing are skewed to the American context (Brayne, 2017; Ferguson, 2017, 2012; Pearsall, 2010). Those scholars that have engaged with empirical research on the use of data systems by police in Europe primarily focus on predictive location policing (Egbert, 2019; Egbert and Leese, 2021; Kaufmann et al., 2019) or theorise about the legal regimes that govern it (Van Brakel, 2020a; 2021a; Kindt, 2020). Yet the scale and scope of technologies developed and deployed within Europe remains unknown. Therefore, to answer my question on the nature of data-driven policing and its relation to police power, I first need to understand what is happening. My empirical research, thus, starts with a mapping study, which is based on a deep reading of grey literature and twelve semi-structured expert interviews in Belgium, Brussels, the Netherlands, and the UK. The expert interviews were conducted with police officers, regulators, members of police oversight structures, and technologists. The sample of grey literature consisted of police, government, civil society, and media websites and reports, as well as a study of data protection laws and government investment portals.

In this mapping chapter, I will outline developments of data-driven policing in Europe. First, I will start with a brief introduction of the European policing context to explain the geographic focus of my mapping and explore the distinct organisational structures that shape the police turn to data. Second, I will explore the specific data-driven policing investment trends I uncovered: augmenting databases, optimising operational support, real-time policing technologies, and predictive policing. This chapter does not intend to give an exhaustive overview of tools that are being developed, bought, and tested but rather approach them as events that offer insight into the changing landscape of policing. I will conclude this section by discussing which investment trends could best serve as my case studies. In the final part of this chapter, I contextualise the police interest in data systems by exploring broader social developments that influence the turn to data-driven policing. Here, I draw on data from all my interviews as experts, practitioners, and civic actors continuously referenced social dynamics in relation to the emergence of these data systems.

4.1 A European policing context

In its essence, the function of a European police force is to be the enforcing arm of the state, with the mandate to maintain public order, ensure safety and security, and prevent and investigate crimes (Bayley and Shearing, 1996). While this top-level mandate applies across police forces we cannot speak of one European police force that operates as a single unit; rather, each country has unique policing structures and political and public mandates. The sheer size and diverse nature of European police forces pose the challenge of where to start with a mapping on data-driven policing. In this first section, I will present my findings that informed the choice to focus this mapping study on the developments in Brussels, Belgium, the Netherlands, and the UK. After this, I will briefly outline each context, specifically the police organisational structures, resources, and cultures, as these impact how police invest in and engage with data-driven policing.

Grey literature informed this thesis that the police in, among others, Belgium, Denmark, Germany, the Netherlands, and the UK are actively experimenting with data-driven policing. These insights are affirmed in the interview with a Schengen evaluator,[4] who performs police data protection assessment across Europe and, as such, has a bird's-eye view of what is happening:

> Let's say that countries that are in the parts of Europe that are more to the north can afford serious investments in IT policing and the countries that are more to the south and are poorer simply cannot afford it. (Schengen evaluator)

Even though these five northern countries are investing in data-driven policing, several experts have noted that there are distinct cultures to how police approach the introduction of new technology. More specifically, the Dutch, Flemish, and Anglo-Saxons share a 'just-do-it' policing culture that favours a more agile, quick, and dirty approach to problem-solving that extends into the turn to data-driven policing. As a senior police expert put it:

> The Dutch are sometimes more influenced by the Anglo-Saxon way of let's just do it and see where we end up. For sure, there are local police units on the Dutch or the Flemish side that identify with that culture of 'yes, let's solve that problem', so we will just do or buy that ourselves, or we will just programme that quickly. (Belgium police manager)

[4] 'The Schengen evaluation and monitoring mechanism monitors the implementation of the Schengen acquis – common set of Schengen rules that apply to all Member States. The purpose of the mechanism is to ensure an effective, consistent, timely and transparent application of Schengen rules by Schengen Member States' (European Commission, 2021)

Another expert described their approach to technology as, ‘at the same time it helps if you just do it yourself. Amsterdam is known for its stubbornness, that we will organise and do it ourselves’ (former team lead). This local resistance connects the push to standardise data practices across the entire Dutch police with the perception that this will slow down their technology development efforts and hinder their ability to get things done. Being stubborn and ‘just doing it’ is perceived to more quickly solve specific policing problems. It is worth noting that civic actors in the UK observe another organisational dynamic, that of individual career incentives, to the police culture on data-driven policing:

> There are career incentives for people within the police that want to be change-makers by bringing some of this in as well. So there is a bunch of incentives and very few it seems obstacles towards using this kind of stuff. (director of privacy NGO)

The perception that individual police officers are rewarded with a promotion for managing a technology project resonates with my observations in UK police meetings and points to a local dynamic that creates individual incentives to experiment with data-driven policing. These insights from grey literature and interviews, where they are experimenting and have a shared culture, informed my choice to narrow ‘Europe’ down to the context of Belgium, Dutch, and British contexts. In addition, several experts argue that we have to situate the police use of technology in these countries[5] within the larger European structures that enable and constrain it. The European Union has several roles, it sets the data protection standards for the police through the LED (EU Directive 2016/680), is an oversight body through the Schengen evaluations, is continuously discussing the expansion of Europol, and is actively investing in the development of data-driven policing tools through its funding programme, Horizon 2020. These observations informed my choice to also pay attention to the developments in Brussels.

The just-do-it culture ties these geographies together, but the organisational structures, resources, and cultures that determine how the police invest in and engage with data-driven policing are distinct. Therefore, I will now briefly outline the policing structures in Belgium, the Netherlands, and the UK. In Belgium, the policing structure is complex. It consists of two levels, the federal and local police, which are autonomous from but interdependent on each other and together form the integrated Belgium police. There are 185 local police units, which are responsible for local police activities such as traffic checks and control of public order. These units fall under the responsibility of the local authorities within the Flemish or Walloon police zones. The federal police is responsible

[5]At the time of my mapping study, the UK had not yet withdrawn from the European Union. This took place at the end of my data collection process

for administrative police orders, judicial police missions, and supralocal police missions, and they fall under the responsibility of the federal government (Vlaanderen, 2021; VVSG, 2021). In this structure, the federal police 'is managing the information support, so the IT, for everyone' (police manager). There is friction in this centralised approach to managing data infrastructures, as those local units that have the resources and interest in data actively experiment with a range of data-driven policing functions. Especially the larger zones in the Flemish part of Belgium, like Antwerp, are pioneering with 'all kinds of systems and with all kinds of IT applications, where the federal police is playing catch-up and has to incorporate these project within their operation' (oversight actor). Thus, the Belgium police structure is both centralised and decentralised, and innovation and information management is officially part of the federal police, but at times, it is initiated by the larger and well-resourced police zones.

In 2015, the Dutch police moved from a decentralised structure, which consisted of twenty-five police forces and one central police force, to one national police. This reorganisation was aimed to reduce police bureaucracy, decrease overhead costs, and centralise IT systems and support. In the current structure, the national police consists of ten regional units and one central unit that are all under the direct responsibility of the minister of justice and security (De Roo, 2016; Government of the Netherlands, 2011; Koning, 2015; Rijksoverheid, 2010). The shift to one national police was believed to advance long-term organisational change on, among others, the 2003 police strategy on intelligence-led policing (Inspectie Openbare Orde en Veiligheid, 2008). In this strategy, modelled after the UK approach, the Dutch police have set the goal to become more information-driven (Den Hengst-Bruggeling, 2013), where police data should be made more accessible and available throughout the organisation for investigation, enforcement, and emergency purposes. As a result, the aspiration to use data for policing has become embedded with the organisation, or as one expert put it:

> The budget is no longer the primary problem, on the hiring side or for the people. We have enough capacity, you see a lot of data scientists coming into the police, who land in different places in the organisation. (Dutch team lead)

Yet at the same time, an oversight actor observed that the police is a large operational organisation: 'at the end of the day, these are implementing organisations, and these units are always far smaller than you think' (Dutch parliamentarian oversight). Thus, while the national police employ over sixty thousand people and have anchored intelligence-led policing within its operations, in reality, only a small part of the organisation is working on and with data-driven policing.

The UK policing structure is again very different from that of Belgium and the Netherlands. It is a federated structure in which England and Wales do not have a national police service, but forty-three individual police forces that are responsible for a specific geographic area (Politics.co.uk, 2021; Varghese, 2010). The Home Office, as the central authority, is responsible but not in charge of the police. An expert explained: 'the UK is unusual because there is no kind of national policing body. It is a federated system where basically every police force is independently and separately run and lots of them kind of acquire surveillance tools and technologies' (oversight expert). While this structure has given rise to forty-three different approaches to data-driven policing, all forces have been confronted with two major shifts: the introduction of the publicly elected Office of the Police and Crime Commissioner (PCC) and the significant austerity measures that have taken place in the last ten years. The establishment of the PCC office was an attempt to decentralise and localise oversight and accountability mechanisms in all areas except for London. They are responsible for 'appointing, and if necessary remove, the chief constable; they set the budget and the council tax precept; and set local policing priorities' (Berman et al., 2012, p. 2). This federated approach to oversight, as such, imposes different standards and requirements on the use of data-driven policing. What sets the UK even further apart from Belgium and the Dutch context is that police forces have been confronted with significant austerity measures that directly and indirectly impact their operations:

> Police have had their budget slashed by 25 per cent in the last ten years. (UK oversight expert)

> There has been a high-level retrenchment in the public services. Policing has often been left exposed to fill the voids that are left by other public services. (UK police adviser)

Austerity measures on other social and public services has confronted police officers on the street with an increased number of mental health issues, missing patients, and suicide attempts. The expert notes that 'because policing has a very broad scope it is very hard for police to realistically say no to things, to say this is not what we do' (UK police adviser). In a context of austerity in which police feel they need to step up and fill the void left by the retreating welfare state, data-driven policing is seen as a solution to do more with fewer resources, which speaks to the managerial logic of efficiency and effectiveness.

In effect, briefly outlining these three contexts shows that, when we approach data-driven policing as a social process, which this thesis does, its emergence is shaped by both the technology and its

surroundings. Here, the organisational budgets, cultures, and structures determine the extent to which data-driven policing is manifesting within the police force, the meaning that is ascribed to it, and if whether introduction is creating new points of conflict within the organisation. In the next sections, I will outline the distinct data-driven policing functions I found in my mapping research. I use the term *function* to describe specific data affordances (Gibson, 1979; Hutchby, 2001) that are ascribed to a suite of data-driven policing technologies.

4.2 Investment trends

> Data-driven policing is very much in its infancy. (Dutch team lead)

> I see that the police are one of the leading organisations when it comes to using data to carry out their duties and goals. (Dutch parliamentarian oversight)

At first glance, these observations might seem contradictory; however, they point to the notion that the turn to data-driven policing within Europe is a relatively nascent practice, but that, in the context of the state, the police is the public institution that is at the forefront of it. These observations could also explain my findings that, while the police are experimenting with a broad range of tools and technologies, some have already been halted and others might never become embedded within their operations (Jansen, 2018; Williams and Kind, 2019). In this section, I will structure my findings on what is happening within the context of Europe along the lines of four investment trends: augmenting databases, optimising operational support, real-time policing technologies, and predictive policing. This will allow me to account for the interest in specific data-driven policing functions without running the risk of this chapter becoming obsolete when individual tools are no longer in use. Thus, this section will not try to provide an exhaustive list of tools that are being developed, bought, and used, but rather use examples to demonstrate specific trends. At the end of this section, I will explore how this mapping informed the focus of my case studies.

4.2.1 Augmenting of databases

The first trend I will discuss is investing in police databases. The collection and storage of crime data have historically been an integral part of policing, where data collected through victim statements, interrogation, observation, investigation, and surveillance techniques were stored in paper files or in what are now called legacy computer systems. To make this data more accessible for day-to-day operations, increase its interoperability, and allow for more novel computational processes, police forces are investing in updating, merging, and enhancing their databases (Bastos

and Curtin, 2020; De Hert and Gutwirth, 2006). Here, I will explore how this trend manifests within the different contexts.

As previously mentioned, the federated policing structure in the UK has produced decentralised IT systems that operate in isolation from each other. To overcome this historically grown database fragmentation, the Home Office is spearheading the National Law Enforcement Data Programme. Through the sub-project the National Law Enforcement Data Service (NLEDS), they aim to merge data from three distinct databases: the federal Police National Computer (PNC)[6], the Police National Database (PND)[7] and Automatic Number Plate Recognition (ANPR) records[8] into one centralised system (Couchman, 2019; Gorcsosova, 2016; Murdock, 2016). To this end, the Home Office awarded £12,000,000 to IBM to assist with the transformation of the existing systems into the NLEDS (Home Office, 2016). In a context of austerity, the Home Office allocates significant resources to augmenting police data infrastructures from the idea that increased access and interoperability will unlock insights that are currently trapped within data silos:

> So this is one big project and this is going to really change the UK landscape, because it is going to provide a platform to knit together all of the different law enforcement bodies into kind of one accessible data set. That's the first phase of NLEDS. The second phase of NLEDS is then to apply an analytics layer on top. (UK oversight expert)

In this sense, investing in accessibility and interoperability is not an end in itself but rather a necessary precondition to ensure future integration and utilisation of a wide range of data-driven capabilities.

A similar trend is observed in Brussels, where new databases are created and old ones are upgraded to facilitate data exchange between police and automate certain identification processes. Several European Member States lobbied for the creation of European Police Registration Information System to overcome what they felt was the too narrowly defined mandate of Europol (Focant et al., 2012). This new database, also described as an index, should facilitate the exchange of police records on criminal activities between the different European police forces through a search interface that provides a quick overview of whether and where a police record about a person exists. After a hit in the system, the querying Member State police force can go through the proper judicial

[6]The PNC is a text-only computer that stores over ten million records about arrest, custody records, and information on suspects, including links to biometric data, fingerprints, and DNA.
[7]The PND stores 'soft intelligence' about allegations and investigations that did not result in an arrest (Murdock, 2016).
[8]The national ANPR database receives around fifty million ANPR 'reads' a day (Police.uk, 2019; Surveillance Camera Commissioner, 2016, p. 23).

channels to obtain access to the information (Jones, 2011). This database is directed at optimising an existing practice by facilitating knowledge exchange between police forces. In addition, Europe continues to invest in the Schengen Information System (SIS). Mostly known as the European Union's border information management system, SIS is in fact a database that contains information on criminal activity, immigration violations, and various objects and missing persons, and it is accessible to European police forces (Jansen, 2017). Investment in SIS was directed at augmenting the database to allow for automated fingerprint recognition. An expert made the points that

> since May, I think, or since March, the fingerprints in the SIS are searchable. So it is the second database with searchable fingerprints. It was introduced with migrants, with asylum seekers, and now EURODAC will receive facial recognition as the very first database. You also have faces in the SIS but these are still not searchable. (European oversight expert)

> As such, the European Union is investing in its data infrastructures from the perspective that it will facilitate cooperation between the different Member State police forces and unlock insights that are currently stored in databases but are not accessible to the human eye. (European oversight expert)

While I did not find a similar investment in the Netherlands, a privacy audit by the national police showed that there are thirty-six critical data infrastructures in which they register, among others, police contacts, arrests, and crime reports, store licence plate numbers, process fingerprints, and exchange information, and none of these are in compliance with their own privacy policies nor the law (Nationale politie, 2020; Zenger, 2020). There is a long-term plan to ensure that their systems will comply in the future. In the Belgium context, police are investing in their database architecture through the iPolice programme, which will be discussed in the optimising operational support trend below. These findings reveal that, on all levels of policing, in a local, national, or European context, investments are made in improving, expanding, and augmenting police databases from the belief that it will disclose insights that are currently locked within existing data silos. This is done as a prerequisite for future data-driven policing practices, to increase coordination between police forces, and to ensure compliance with data protection standards.

4.2.2 Optimising operational support

The second trend I will discuss is that of investing in technologies that support the operational processes of police. These investments range from improving the case management system for the

documentation of criminal offences and the creation of dashboards to visualise police data to mobile apps that allow police to access police databases, report incidents, and in some cases check biometric features while out on the streets (Gemeente Den Haag, 2020; HLN, 2016; Richard Vis, 2020). Investments in these tools are aimed at removing the need for front-line officers to go into the office and to speed up certain processes, which should allow officers to spend more time on the streets.

The Dutch national police developed the MEOS (Mobile Effective on Street) app, which allows police officers on the street to directly access police information systems through their mobile phones. The rationale is that police officers can do specific data processes, such as scanning identification documents and licence plates to determine the identity of an individual or a vehicle, and register criminal offences in the field, and they do not need to come back to the police station to do their basic police work. In 2017, 35,000 officers were working with the app throughout the country (Algemene Rekenkamer, 2016; Nationale politie; 2017). Thus, MEOS is believed to increase police efficiency, as they can implement a range of different tasks without having to go back to the office. A team lead explains:

> Things like MEOS help. Let's take the example of the phone: you will no longer need to ask and verify information on your radio, but you can already check things yourself from biometric features, drivers licence, and that sort of stuff – that is amazing of course. Another benefit: it increases the motivation of our colleagues, to know that they are working for a professional organisation.

Beyond facilitating access to data systems on the street, it is clear that advancing these technologies are also perceived to signal to front-line staff that they are part of a professional organisation that is constantly evolving and keeping up with the times.

In Belgium, the federal police are undertaking a large-scale digital transformation programme, under the name iPolice. After the Brussels terror attacks in 2016, the Belgian government has allocated an additional 110 million euros to improve data exchange between the different police entities, support officers in the field, and increase their overall effectiveness, which resulted in the project iPolice (HLN, 2016). This ambitious case management system aims to overhaul and improve the information architecture and operational processes of the Belgium police. The system is designed to be modular so that, when needed, more novel technical capabilities can be added at a later date. At the time of the interview, iPolice was expected to be operational in 2020 and would

include a mobile app, similar to MEOS, but with the additional feature that investigators can push a data collection request to front-line staff (HLN, 2016).

In the UK, media have reported on the use of mobile fingerprint scanners to check people's identity on the streets (Richard Vis, 2020), and an expert drew attention to investments in 'kiosks' that should allow police to move part of the forensic data collection process onto the streets:

> So a lot of projects have been around what is termed kiosks, self-serving computer workstations that allow phones and other devices to be downloaded, without them having to go to a lab and go through a rigorous forensic investigation, which has often very long backlogs in place. (policy adviser)

Again, the introduction of these apps and kiosks indicates that police see value in investing in technologies that will allow them to move part of their existing organisational process to the streets. These findings reveal that, in part, the police believe they become more effective if and when their front-line officers spend less time behind a computer in a police station and more time on the streets. Investments in this area are thus directed at optimising existing police practice, which is being visibly present on the street.

4.2.3 Real-time recognition systems

The third trend is that of investing in technology that allows for the real-time identification of individuals and objects. In this section, I will discuss Automated Number Plate Recognition (ANPR) and biometric recognition, in other words, (automated) facial-recognition and voice-identification systems. Here, I differentiate between biometric recognition and the more established forms of police identification, such as fingerprints or DNA identification, as these more novel forms of recognition can automatically collect data in the 'wild' from security and social media data infrastructures without the individuals' knowledge or consent, and they can retrospectively or in real time be compared to an existing database (Gates, 2011; Surveillance Camera Commissioner, 2016).

ANPR is a widely adopted real-time identification system in Western Europe that collects immense volumes of data to enforce traffic violations and monitor movements of vehicles across highways and cities' access roads. In grey literature, I found that, in the UK, ANPR systems had 'approximately 8,500 cameras in use capable of capturing 35 million and 40 million "reads" a day and storing upwards of 30 billion "reads" a year' (Surveillance Camera Commissioner, 2016, p. 23). In the Netherlands, the police has about 450 road cameras and 150 mobile units at their disposal,

and the data collected through it can be stored for up to twenty-eight days (Privacy First, 2019, 2018; Nationale politie, 2019a). Despite the wide use of ANPR and the volume of data collected through it, this function was only mentioned in passing by experts. This points to a dynamic in which the visible use of surveillance tools normalises its existence over time. Science and Technology Studies scholar Star (1999) has famously argued that it is in moments of breakdown that infrastructures become visible. In the context of ANPR systems, moments of breakdown materialise on the personal and political, the moment when one is confronted with a ticket or when changes to the scope and nature of the technology again become part of the public debate (NOS Nieuws, 2017).

Facial recognition, and specifically live facial recognition, is not widely adopted but emerges in the public view through experiments and pilot programmes. British police have been the most visible police force piloting this function on the streets. The Met tested facial recognition for public order policing and have since halted its use. The test was done by deploying mobile camera units at the Notting Hill Carnival and on the city streets for the identification of persons of interest (Big Brother Watch, 2018; London Policing Ethics Panel, 2018). In Cardiff, the South Wales Police received £2.6 million from the Police Transformation Fund to lead the testing and deployment of facial recognition. ‘South Wales Police has admitted it has used automatic facial-recognition technology to target petty criminals, such as ticket touts and pickpockets outside football matches, but they have also used it on peaceful protesters’ (Liberty, 2018). This system is still in use, and on their website, they keep a record of the deployments that have happened in the past (South Wales Police, 2021). In the Netherlands, the forensic department of the Dutch national police uses Catch, a face comparison tool, in the context of specific investigations, and they are experimenting with the development of a privacy-by-design facial-recognition system at the Johan Cruijff Arena, Amsterdam’s football stadium (Amsterdam, 2021). In Belgium, the federal police tested a facial-recognition system on the airport near Brussels, which has since been halted by the data protection authorities (COC, 2019).

A more novel form of biometric recognition is voice identification, which is also referred to as speaker recognition. A survey conducted by Interpol in 2016 found that twenty-two law enforcement agencies in Europe had speaker identification capabilities (Morrison et al., 2016). The report is less clear on how these capabilities are deployed and where, but it suggests they are primarily done by forensic practitioners who make make quantitative measurements on speech recordings ‘who carefully select and prepare recordings which are then analysed using signal-processing algorithms (Morrison et al., 2016, p.94). One expert (UK oversight) mentions the

Speaker Identification Integrated Project (SIIP), a European research project that is developing a voice system that will allow for the automatic identification of individuals on the basis of their voice. Like facial recognition, automated voice identification is believed to unlock insights that are currently stored in police databases. In the context of iPolice, the Belgium police manager explains:

> If we buy the option voice to text, yes, then we can automatically convert all tapped material to text. You can save a lot of time with this. More importantly, voices are very unique, so that offers new opportunities for the police. (Belgium federal police manager)

In this sense, biometric recognition technologies allow police to reduce the amount of time it normally takes to transcribe voice evidence, and it is the versatile nature of biometric recognition technology that speaks to a social-technical imaginary of how it can address several operational needs and challenges in the near future. These findings reveal that police across jurisdictions are investing in real-time recognition systems, which range from ANPR to facial recognition and voice identification. In addition, when it comes to biometric recognition technologies, this mapping found that the projects are not widely adopted but are rather more experimental and aspirational in nature.

4.2.4 Predictive policing

The fourth trend that can be distinguished is the investment in predictive policing technologies, in which we can distinguish between tools that predict the possible location of a criminal offence or the individuals who are most likely to commit or become a victim of crime in the near future. Investments in these tools are aimed at changing how and when the police intervene, where they can deter future criminal offending by being at the right place at the right time or by actively monitoring and interfering in the lives of people at risk of committing a crime in the future (Brayne, 2017; Egbert and Leese, 2021; Ferguson, 2017). In this section, I will discuss these predictive technologies as separate data-driven policing functions.

Predictive location policing, also known as hotspot policing, might be the most tested and ephemeral function I found in my mapping study. There are a wide range of products being developed and deployed that aim to predict the locations of high-impact crime (HIC), such as robbery, theft, and burglary, and police forces are equally stopping its use after initial trials. In the UK, Liberty found that, in 2018, thirteen police forces used, trailed, or developed one or more predictive policing programmes (Couchman, 2019, p. 45). The Met was found to have trialled the commercial products PredPol, Azevea, and Palantir, while at the same time developing their own in-

house system. Since these initial trials, a cost-benefit analysis resulted in halting the use of the commercial tools (Beckford, 2018; Couchman, 2019). Kent Police spoke publicly about their turn to PredPol in the fight against acquisitive crime but have since halted its deployment in 2018 (Chowdhury, 2018).

There is little known about the use of predictive policing tools in Belgium. The scholar Van Brakel (2020b) reflects that police probably are not deploying much yet, but the lack of publicly available information on the topic makes it difficult to assess whether and where it is being used. It is important to note that a Belgium police coordinator observed at the end of our interview that the crime rates in Belgium and the Netherlands are too low for predictive analytics to work. Still, the Dutch police, who invested in the development of a predictive location tool called the Crime Anticipation System (CAS) (Williams, 2014), did not halt its use after disappointing results but rather introduced it within the national police (Korpsleiding, 2021). A policing expert reflected on CAS:

> Now, CAS mainly supports the information organisation. You see how these forecasting models are shown in the police cars in America. That they drive into a neighbourhood and that a window pops up with a note that there is an increased risk of burglary here, pay attention to that and that. That sounds very nice, but it just doesn't work for us. We tried it. (team lead)

That the predictive policing tool did not work as it was intended to did not mean that the tool itself had to be discarded. It was repurposed and became another piece of intelligence that is provided to police managers on a biweekly basis to inform decision-making on staffing rosters. This suggests that, when police allocate resources to the development of a specific technology and it is supported by internal advocates, it will one way or another become embedded within the organisation.

Predictive identification, also known as data-driven risk scoring, identifies and ranks individuals according to the likelihood that they will engage in criminal or violent activity in the near future. Its use is primarily reported on in the UK and the Netherlands. Probably the best-documented and most publicly criticised European predictive identification programme is the Gang Matrix, deployed by the Met, which aimed to identify potential gang members and score them according to the risk they pose to society. After public critique that revealed the discriminatory nature of the Gang Matrix (Amnesty International, 2018; Scott, 2018), the Information Commissioner's Office (ICO) and the mayor of London launched an investigation that temporarily halted the programme (ICO, 2018;

Mayor of London, 2018). Other predictive identification programmes are the Integrated Offender Management model of the West Midlands Police, who also lead on a nationwide project called the National Data Analytics Solution (NDAS) project. NDAS is funded through the Police Transformation Fund of the Home Office (Home Office, 2018) and is developed in collaboration with eight other police forces. In both cases, ML and statistical modelling are applied to categorise and score the risk of someone committing or becoming a victim of a number of predefined crime priorities, such as knife crime and modern slavery (Baraniuk, 2018). In the Dutch context, there are a number of different data-driven risk-scoring programmes, which police refer to as a person-based approach. The most notable are the Top 600 and Top 400 by the Amsterdam municipality, which selects individuals for a care and control approach on the basis of police, public prosecutor, and social services data. Once identified, the idea is to holistically intervene and deter repeat offenders and potential criminals from engaging in future criminal activity (Abraham et al., 2011; Openbaar Ministerie, 2019). Another tool, ProKid-12, is being developed in the east of the Netherlands and aims to predict which children under the age of twelve are showing concerning behaviour on the basis of police data (Abraham et al., 2011; Wientjes et al., 2017). Police do not act upon the output but view it as a referral mechanism to flag their concerns to other public authorities.

I will conclude this section on investment trends with the observation that even though the turn to data-driven policing is a relatively new and nascent practice, police in Belgium, the Netherlands, and the UK are actively investing in it. More specifically, these countries are investing in augmenting databases, optimising operational support, in real-time policing technologies, and in predictive policing to enhance the day-to-day operations of their police. These findings not only contribute to a broader understanding of how data-driven policing is emerging within the context of Europe but also inform the selection of my case studies. When I look at the investments that are made into the augmenting of police databases, it is rooted in the belief that it is a fundamental infrastructure needed for policing that will in turn also enable other data-driven practices. However, these infrastructural developments are rather diffuse and far removed from the day-to-day of policing, which makes it difficult to isolate this function for a specific case study. The investments into operational support tools are perceived by practitioners as a way to optimise existing operational processes for police on the street. Apps and kiosks, as a mobile extension of police data infrastructures that, up till now, have been tied to a physical location, would be an exemplar research project for those studying the impact of technology on front-line staff but are less well suited for a case study into how data is mediating police understanding of crime and police power. ANPR systems have become so normalised that they are only mentioned when police intend to layer this system on top of or integrate it within other data-driven policing functions. Predictive

location policing is a much-studied data-driven policing function (see chapter 2), but my mapping revealed that these projects are ephemeral in nature and different projects have since seen their introduction halted. However, the praxis of data-driven risk scoring and biometric recognition are often stand-alone projects that are introduced with the aim to bring something new to how the police operate. Thus, these two functions provide clear examples of how data is changing the priorities, needs, and action repertoire of the police, and these examples are well suited to inform my analysis on the relationship between datafication and police power.

4.3 External dynamics

The central question of the thesis concerns itself with the nature of data-driven policing and the relationship between datafication and police power. Here, I argued that this first requires an understanding of what is actually happening in Europe. To this end, I engaged in a mapping study that explored where police forces in Europe are investing in data-driven policing and clustered these developments along the lines of four trends: the augmenting of databases, optimising operational support, real-time policing technologies, and predictive policing in Brussels, Belgium, the Netherlands, and the UK. My mapping, as such, informs this thesis on the specific investment trends but less so on how these developments are shaped by the environment police operate in. In the final section, I will account for their context, more specifically the extent to which these developments are dependent on the changing nature of crime, the social and political context within which they operate, and the oversight bodies that govern their turn to data. I do so because social scientists have theorised that the authority of the police and the justification of their actors is dependent on the authorisation of the state and other external stakeholders. As such, I'm interested in understanding how external stakeholders influence the turn to data-driven policing. I draw on the observations and reflections of all my interviews to conclude that the police are strategic agents who negotiate the internal and external expectations placed upon the institutions, in which the external environment primarily creates an enabling environment the use data systems to materialise.

4.3.1 Changing nature of crime

The mapping of investment trends reveals that, despite the range of applications, all these data solutions are aimed at improving the existing operations of the police. Take, for example, investments in databases, which are aimed at unlocking insights that are currently stored in the data silos, or how predictive policing is directed at improving responses to HIC and acquisitive crime offences. This suggests that the understanding of crime and police crime priorities remain the same,

and police investments into data systems are primarily directed at changing how crime insights are generated. In this section, I will explore the nature of crime as observed by the experts vis-à-vis police crime priorities to conclude that the focus of the police, and their use of data-driven policing, is the result of at times competing and conflicting internal and external dynamics – a negotiation that is characterised by a general reluctance from the side of the institution and its practitioners to change priorities along shifting crime phenomena, but certain high-profile incidents demand a response from them.

The Belgium, Dutch, and British crime context is characterised by a significant decrease in HIC and acquisitive crime rates and a rise in other criminal offences, such as cybercrime, violent crime, and terrorist attacks. In Belgium, an expert observes,

> here in Brussels criminality has decreased by 40 to 50 per cent in the past ten years, so what do want to direct our attention towards? ‘cause there is a large dark number on cybercrime, incidents impacting individuals, but also the security of a country and that of companies. We spend too little time on this. (police manager)

The mapping did not reveal any significant investments into technologies that would allow police to more actively monitor and fight cybercrime. This in part can be attributed to the sample of experts interviewed and grey literature studied for the mapping and in part to the international and specialised nature of investigating cybercrime. Take, for example, the Dutch police, who define cybercrime as a cross-border phenomenon and has anchored the investigatory powers within a specialised team (Nationale politie, 2021): team high-tech crime (police), the public prosecutor’s office, and a national cyber security centre (Ministerie van Algemene Zaken, 2011). However, in part, this is also the result of organisational dynamics that constrain a shift in police priorities. The police manager continues to explain that ‘it is difficult, ‘cause they have their full-time equivalants and they would not want to give these up’, indicating that the historically determined organisational structure and budget allocation limit the agility of the police to adapt to new criminal offences that might require a different approach, skill set, or team composition. A similar dynamic can be witnessed in the Netherlands, where burglary, theft, and robbery is on the decline (Nationale politie, 2019b), and other criminal offences, such as cybercrime and organised crime, are on the rise. Yet a former team lead of the Dutch police explains:

> Our focus is on the street; our focus is on HIC, burglary, theft, and robbery. Once we control this, then we can look at what else is needed. Is there a security plan, making

> deals with prosecutors and mayors? That is one thing, but you could start to look at other things, like cyber, digital world.

Thus, the police ability and interest to change is dependent on the police belief that they have a handle on historically determined crime priorities and those criminal offences that get prioritised by external stakeholders. Here, the expert refers to the 'triangle', a consultative body between the representatives of the police, public prosecutor, and local authorities within a specific geographical area who discuss and coordinate interventions around public safety issues.

A similar trend can be witnessed in the UK: 'over the last twenty years, there is a significant reduction in traditional volume crime, which was predominantly acquisitive crime, and a significant increase in other forms of crime, particularly violent crime' (police adviser). However, he observes that these new types of criminal offences are not recorded properly in the two authoritative crime data sources: the police crime records and the Crime Survey for England & Wales.[9] As such, the size and impact of violent crime and cybercrime remain unknown. Still, historic responses to high-profile violent and immoral offences show that times of crisis are a catalyst for change.

> If a terror attack occurs, there will be an enormous injection, an impulse to the development side [of data-driven policing]. If only to keep the troops calm, calm, calm. (Dutch team lead)

> Yes, this part is because of the terror. First, we had Dutroux, who enabled us to create a national database, which we were able to run statistics and that phenomenon became a little bit more accepted. The 100 million [euros] was made available after the terrorist attacks. The plan was already on paper but it was never going to get money. (Belgium programme manager)

In times of crisis – whether the Dutroux affair, the kidnapping, torture, abuse, and murder of young girls in Belgium in the 1990s, or the violent terror attacks that have confronted Europe in recent years – technology projects that would under normal circumstances not have been a priority in the allocation of police resources get accelerated.

These developments show how, even though society is confronted with changing criminal offences, the police is a strategic agent that primarily responds to criminal offences that are in its wheelhouse, the ones that are most visible on the streets or have political momentum behind them. The

[9]The Crime Survey for England & Wales is a victim perception study of crime conducted by the Office for National Statistics.

institution is less able to respond to emerging criminal offences that are outside their realm of everyday practice and do not have sufficient political momentum. As such, data-driven policing supports the police to optimise for existing crime priorities and new ones that emerge from the interplay between internal and external dynamics.

4.3.1 The social and political context

In chapter 2, I drew on social scientists to argue that the police are the most visible agent of the state, who are deeply implicated in enforcing the political orders through the threat of punishment (Bourdieu, 1991; Jackson and Bradford, 2009; Loader, 1997). Here, the authority of the police and the legitimacy of its practices are shaped by and dependent on the authorisation of the state and other external stakeholders. This requires a mapping of what is happening in the context of Europe to account for the social and political context that enable or constrain the turn to data-driven policing. I will draw on all my interviews, as experts, practitioners, and civic actors continuously refer to social and political dynamics to explain and or justify the broader turn to data-driven policing. In this section, I will start by exploring the political and social pressures experienced by the police and then move on to the financial incentives that are accelerating its use. I will conclude that external stakeholders are creating an enabling environment for the police to use data and at times even actively endorse it.

The police feel they have limited autonomy in setting their priorities, and the turn to data-driven policing, as such, can only materialise within the boundaries set by political and public debate. One practitioner noted that,

> we can say what we want or don't want [with data-driven policing] as a police [force], but we have a limited mandate to determine this, of course. That's what politics actually, society, has to indicate. (Dutch police manager)

At the same time, it is clear that the external environment within which the police operate is not immune to the ideology of data. The interest in data, 'it's not just the police, it's also politics that plays a role in this, the pressure from politics' (Belgium data protection actor). Politicians put pressure on the police to turn to data in the belief that this will allow them to better manage and control excessive and unwanted behaviour in society. As one police manager said, 'after an event, politicians are demanding a safety utopia, where we want to ban every risk'. Another practitioner described it as 'the belief that we can control society and people' (police adviser); this was articulated by a data protection actor:

> Personally, I see a broader societal desire to eliminate risk. So it is not just about wanting people to follow the rules and punish them the moment they've crossed the line. But there is a desire to make sure that people can't cross the rules to begin with, so that's actually the risk-free society. (data protection actor)

In that sense, police practitioners feel that the social and political desire to manage and pre-empt behaviour that is considered abnormal and unwanted is placing normative expectations on police, where the use of predictive analytics will allow them to adhere to them.

The social and political desire to eliminate risk does not imply that political discourse merely endorses its use – 'the parliament asks a lot of questions related to privacy' (Belgium police manager) – but rather foregrounds a broader societal shift in which data is seen as an authority to control and manage crime. As such, the turn to data-driven policing is not an isolated event but is part of a larger social transformation.

> Obviously, they've got no good example being set in government because, especially during the pandemic, the government is doing exactly the same thing. So, yes, the overall culture around the adoption of predictive analytics, AI, automation technologies that use personal data is very enabling. (UK privacy expert)

The broader turn to predictive analytics is particularly important in a context in which the police have been confronted with significant austerity measures and there is an absence of political leadership on policing, as this further reduces the threshold to use of data-driven policing. As one expert observed, 'there has been a distinct absence of any strategy around crime in general over the last decade as well as a conscious political effort to distance the government from the impact of austerity' (UK police adviser), leaving the forty-three UK police forces alone to set priorities and make operational decisions in relative political isolation. At the same time, police are actively encouraged to experiment with data-driven policing functions through short-term funding opportunities offered by the Home Office's Police Transformation Fund.

> The PTF [Police Transformation Fund] has been ill conceived and poorly executed, because what it has done is incentivise police forces to embark on projects often using data-driven capabilities, but without having a clear strategy about what the end goal of the data-driven capabilities is. (UK police adviser)

Similarly, an expert observed that the mushrooming of data-driven policing tools in the UK is the result of political pressure and funding opportunities:

> Partly, there has been a top-down pressure on police forces to adopt more technology. And partly it is a reflection of the availability of funding to support data-driven initiatives through the Police Transformation Fund. (UK police adviser)

While the UK context, with its austerity and political vacuum, is unique, it is worth noting that my mapping found that financial incentives are an accelerator of data-driven policing practices across Europe. The European Commission is stimulating the development of the next suite of data-driven policing tools through its security funding programmes under Horizon 2020 and its predecessor Seventh Framework (FP7) Programme (European Commission, 2018; Leufer and Jansen, 2020; Mazzucato, 2011). And the Belgium federal government made 110 million euros available for the development of iPolice after the Brussels terrorist attacks (HLN, 2016).

These findings reveal that, while the social and political contexts in which police operate continuously changes and police priorities and actions are continuously renegotiated in response to the external demands and expectations placed on them, the state and other external stakeholders are also subjected to the dominant social norms and values that are shifting with the datafication of society. As such, they authorise the use of data by creating an enabling environment for the police to use data and at times even actively endorse and enable this use of data.

4.3.2 Oversight mechanisms

I conclude this section on the European context in which data-driven policing emerges by exploring the role of data protection oversight bodies, as this external stakeholder is perceived to be an important body that supervises and at times restricts the use of data in the context of policing (De Hert and Sajfert, 2018). Throughout the interviews, data protection experts expressed concerns about the collection, storage, and use of unverified data, the meaningful role human decision-making plays in the use of automated systems, and the at times rushed manner in which technologies are bought and tested. While each of these concerns could be an empirical chapter in its own right, what is important for my analysis on the relationship between data-driven policing and police power is how these oversight mechanisms perceive, enable, or constrain the use of data-driven policing. Therefore, this section will outline the observations made by data protection practitioners about their mandate and relationship to the police. I conclude that, contrary to popular

belief, their role is not to prevent or restrict the turn to data-driven policing, but rather ensure that the police comply with the rules set forth by the state.

The oversight practitioners I interviewed think that society, and the police, primarily see them as an entity that constrains the use of data in society – so much so that data protection authorities feel that the police at times dismiss their concerns by positioning them as an actor that is against technological development. As such, they consciously describe their mandate as not being against the use of data but rather ensuring that it happens in accordance with predefined rules and regulations. One data protection authority described:

> The ICO always says we are not at all against technology and innovation, but it has to be done in a way that is consistent with the law. (UK data protection actor)

In response to my question on why the Belgium oversight authorities specifically mentioned that they are not against the use of technology by police in their report on a specific data-driven policing function (COC, 2019), the data protection actors responded:

> We are absolutely not opposed to this, quite the contrary. But it is also necessary to clearly say so from time to time, so this is primarily a message addressed to the police services themselves. (Belgium data protection actor)

In this sense, data protection authorities feel they need to actively signal to police and others that they are not against the use of data. This observation is particularly pertinent considering that data protection practitioners have both an advising and enforcement role, which interviewees also refer to as the ‘carrot-and-stick’ approach. How they straddle these two distinct roles is up to the discretion of the data protection authority. As one data protection officer outlines:

> We try to think along and have many conversations in the field; we have good contacts, also within the police. We don’t want to sit on their lap as a supervisor, and we don’t want to be pushed around either, but we are emphatically looking for the common goal and finding each other in it. (Dutch data protection actor)

The emphasis on good contacts and a common goal speaks to the notion that data protection oversight mechanisms are constructed within the Westphalian view of the nation states (Fraser, 2008). Here, the state is seen as the primary power holder that grants rights to its subjects and harms are negotiated within its structures. Oversight bodies are one mechanism created by the state to

negotiate harms within the structures set forth by them. As such, data protection authorities and the police are both a part of the production and justification of the same political order, and oversight mechanisms that aim to supervise the fair use of data by police do so according to the rules set by the political entity from which they both emerged.

It is worth noting that there is a sense that the requirements for the police to use data are much higher than commercial entities working in the security industry. Here, data protection officers referred to private detective agencies, companies that do background checks on the basis of publicly available information, and other security companies. As one practitioner put it, 'so the police rightly have to jump through hoops, and I think those hoops are really justified, but I think it's not quite fair that other organisations don't have to' (Dutch data protection actor). The turn to data-driven policing is believed to be measured against two standards: the impact of police actions on people and the fairness of zooming in on police actions and not that of commercial security actors in society. This begs the question of when and how do data protection authorities intervene? There are different avenues that inform the actions of data protection authorities, such as complaints, pressure from civic actors, and news reports. A data protection authority practitioner reflects:

> The ICO is not overrun with complaints from the public about the use of AI. We receive very few complaints about its use. It is generally civil libertarians and parliamentarians that question its use. (UK data protection actor)

Yet at the same time, this does not mean no action is taken. As the same practitioner put it: 'just because lots of people are not complaining about it, does not mean that lots of work does not have to be done on it to make it right, in order to make it fair while there are still opportunities for good practice to be embedded in' (UK data protection actor). 'Make it fair' and 'good practice' already presumes an inevitability in the use of data-driven policing, as such, the guiding question is not if but under which conditions these functions can be implemented. The role of the data protection authority is not to prevent the turn to data-driven policing but rather to ensure that the police comply with the principles layout in the LED (EU Directive 2016/680).

I will conclude this section on the relationship between the external environment and the police ability to turn to data-driven policing by arguing that the police are a strategic agent who negotiate a range of expectations that are placed on them, as illustrated through their responses to the changing nature of crime and the normative expectations that are placed on them by the public and politicians. This external expectation becomes especially visible in times of crisis, which can range

from a high-profile violent attack or austerity measures, in which political structures actively promote the use of data-driven policing through discourse and the mobilisation of financial resources. Even oversight mechanisms that emerge from and are implicated in the production of our contemporary political order that privileges the use of predictive analytics to solve predefined social problems. To conclude, the police operate in an overall enabling environment that incentives their turn to data-driven policing.

4.4 Conclusion

What this chapter has done is contextualise the turn to data-driven policing within the context of Belgium, Brussels, the Netherlands, and the UK. My mapping across these different geographies has contributed to an understanding of what is happening in Europe and how the broader turn to data-driven policing is dependent on both the police organisational cultures and structures and the wider political climate within which they operate. I structure these findings along four investment trends: i.e. augmenting databases, optimising operational support, real-time policing technologies, and predictive policing. Where I show that even though data-driven policing is still a nascent practice, police forces are actively developing, buying, and testing a multitude of tools that are believed to optimise their operation. This chapter concludes by approaching data-driven policing as a governance issue that is dependent on the interplay between the police and external stakeholders, in which external actors primarily create an enabling environment and at times actively endorse the turn to data-driven policing.

This mapping chapter further informs the rest of my research in a number of ways. The ephemeral nature of specific data-driven policing technologies, in which police invest in a wide range of tools that most likely never materialise into practice, has influenced my decision to look at functions rather than a specific tool. Here, I build on Egbert's (2019) observation that it is the perceived affordances of specific functions that are changing the practice of police and less so the specific tools that are being developed and tested. Furthermore, the analysis of the investment trends has informed the selection of my case studies. It foregrounded functions that can be considered stand-alone projects that are introduced with the aim to bring something new to how the police operate, which will allow me to study how data is changing the priorities, needs, and action repertoire of the police. As such, I have chosen data-driven risk scoring and biometric recognition as my two case studies, as these two functions are best suited to studying the relationship between data and police power. Finally, the mapping chapter demonstrates the value of looking across and between policing contexts, to distil overarching patterns that speak to the mechanisms and social structures from

which they emerge. Building on these insights, the next two chapters will engage with the actual practice of data-driven risk scoring and digital biometrics across different European countries.

5. The practice of data-driven risk scoring

Data-driven risk-scoring models, which aim to identify those individuals who are most likely to commit or become a victim of a specific crime type in the near future, are essentially statistical models that analyse historic data sets to look for one or more group traits belonging to a specific criminal offence (Harcourt, 2008). While this description explains the basic functioning of data-driven risk scoring, it decontextualises the practice from its social, political, and operational environment that shape and are in turn shaped by these technologies. In this chapter, I will discuss the findings of my first case study, in which I will explore the use, experiences, discourses, and practices around data-driven risk scoring in the Netherlands and the UK, which contributes to answering my two research questions: what is the nature of data-driven policing? And what is the relationship between datafication and police power? This chapter is based on the data I collected through twelve police practitioner interviews, as well as participant observation in police meetings and the study of grey literature. My interview sample consisted of senior police managers, specialised police officers, municipality staff members, and relevant oversight bodies. I choose to focus on researching data as practice, as it allows me to not only observe what is happening but also to engage with the belief systems that inform how the concept of risk is constructed and how it is shaping the policing organisations.

Data-driven risk-scoring model	**Police**
Top 600	Municipality of Amsterdam in cooperation with, among others, the Amsterdam police force
ProKid	Gelderland-Midden police
IOM model	West Midlands Police
Domestic violence ML	Hampshire police

Figure 3: Overview of data-driven risk models studies discussed in this chapter

This chapter is organised as follows. First, I introduce four distinct data-driven risk-scoring models (see figure 3), each one representing a stand-alone project that is introduced with the aim to bring something new to the operation of policing. I want to note that this case study does not intend to offer a comparative analysis between the distinct deployments, as it will lack the breadth and depth needed; instead, it aims to identify broad organisational principles that structure police approaches to technology and offer insights into the social structures that (re)produce them. Second, I will explore to what end risk is constructed. My research shows that practitioners inscribe multiple

meanings to risk, such as a normative measure of behaviour, a justification of police intervention, and to increase coordination between different state actors. As such, while this function is often positioned as a sociotechnical system that attributes a level of riskiness to an individual, it should primarily be understood as a construct that is shaped by and is shaping policing as an organisation, a dynamic that I will refer to as the *internal organisational optimising logic*. To the end, I will outline the new organisational practices that emerge from the turn to data-driven risk scoring.

5.1. Data-driven risk models

This chapter will first explore the practice of data-driven risk scoring by outlining four distinct models that are developed across the Netherlands and the UK: the Top 600, ProKid, the Integrated Offender Management model, and the ML in domestic violence model. Figure 4 provides an overview of the data-driven risk models that will be discussed, their focus, and the criteria on which the riskiness of individuals is constructed. For each of these models, I will describe how they came to be, their origin stories, and how risk is calculated, to conclude that the function of data-driven risk scoring emerges from specific political, organisational, and individual interests and form a top-down approach to crime. This, I argue, requires a deeper understanding of to what end these risks scores are implemented.

5.1.1 Top 600 – the Netherlands

Since 2011, the Top 600 aims to reduce the number and impact of HIC offences in the city of Amsterdam (Gemeente Amsterdam, 2019). The following offences are classified as HIC: robbery, street robbery, burglary, aggravated assault, murder, and assault. The Top 600 is a prime example of the person-oriented approach that is implemented throughout the country and emerged from a distinct political intervention. Politically, there was a sense that the traditional interventions of arrest and punishment were not deterring a small group of prolific HIC offenders, and after a number of high-profile violent offences, there was significant momentum to reduce the number of violent HIC offences.

Name	Focus	Databases	Criteria
Top 600	Identify the most prolific HIC offenders and those most at risk of engaging in offences related to drug crimes and violence in Amsterdam	Police and public prosecutor data	1. Criteria for prolific HIC offenders Police criteria: • have been apprehended as a suspect in a HIC event in the last five years and/or • have been a suspected of a felony in the last two years Justice criteria: have been arraigned before a examining magistrate in the last five years, has had contact with the public prosecutors at least three times and convicted at least once in the last five years. Under twenty-one, if they have come in contact with the public prosecutor at least twice and convicted once in the last five years or after a conviction in a HIC incident. 2. Criteria for drug and violence Police criteria: all individuals under thirty that are in the police registration system and have been arrested in the past five years for selling hard drugs AND a violent crime, threat, or possession of weapons Justice criteria: and who have been sentenced to imprisonment or community service for a crime at least two times in the past five years, of which at least one time in the past three years; • or imprisonment or community service for Section 2 Opium Act, with the exception of Section 2C (having a dealer present without a dealer indication) • or a prison sentence for one of the articles in the Weapons and Ammunition Act • or imprisonment for a serious crime of violence (including attempted)

ProKid-12	Identify children under twelve years old who show concerning behaviour and refer them to relevant public authorities.	Police data	Incidents and activities of a child under twelve registered in police systems, as a perpetrator, victim, or witness of crime, and the incidents and activities related to their home address.
Name	**Focus**	**Databases**	**Criteria**
ProKid-23	Identify those adolescents under twenty-three years old who show increasingly worrying behaviour that indicates a probability of future registration as an offender. This related to general offences, but can be narrowed down to violent offences, vice, offences with violence, or offences without violence.	Police data	Incidents and activities of a young adult, and the incidents and activities of their co-suspects, where there has to be at least one common police registration. Independent variables are • age, • gender, • incidents of the minor, • variation of the incident of the minor, • activity of the minor, • incidents of the co-suspects, • activities of the co-suspects, and • number of co-suspects.
IOM model	Identify acquisitive crime offenders most at risk of re-offending and escalating from low and medium to high harm offences.	Police data	Fifty variables from eight police databases: Crimes (crimes committed), IMS (intelligence), ICIS (custody), PINS (prison notification system), Corvus (intelligence and tasking system), OCG (organised crime group data), OASIS (the event logging system), and DiP (drug intervention programme data). SAS (stop and search) is no longer included in the model.
Domestic violence ML model	Identify those at risk of committing, or being a victim of, gang-related violence in London.	Police data	148 variables from Dash registration and police database

Figure 4: Data-driven risk models

> Seven or eight years ago, that was totally a thing, that was big in the news. We had a couple of fatal robberies of a store and a jeweller, and the number of burglaries and street robberies was pretty high, so we have that category of crimes, from that, administrators said, yes we need to do something with this. (CIO)

The Top 600, a political response by the former mayor of Amsterdam, was designed to identify a group of prolific offenders and structurally intervene in their lives. It aims to prevent crime by assessing the risk of re-offending and by minimising the number and impact of future crimes committed by known offenders through controlling their behaviour and changing their life conditions. In addition, it aims to prevent the influx of younger siblings by minimising the negative impact of the lives of known offenders on their families. This is also known as an integrated care and control approach, which combines activities of surveillance, arrest, and conviction with interventions aimed at addressing the multitude of problems that exist in the life of an individual labelled as risky, such as debt, unemployment, addiction, psychosocial problems, and mental disabilities (Ferguson, 2017; Van der Put et al., 2013). It is important to note that the Top 600 is housed within the city of Amsterdam and is set up as an effort to increase collaboration between a range of public authorities, which include the police, public prosecutor, and the GGD.[10]

To identify who qualifies for the Top 600, police analysts developed a relatively simple data-driven risk-scoring model, which biannually creates an initial shortlist. This risk-scoring model is designed to select the most prolific offenders who reside in Amsterdam from the total number of HIC offenders present in police databases and in public prosecutor data. As the chief information officer (CIO) outlined: ‘We discovered pretty quickly that there are about fifteen thousand people in the police systems who had something to do with a HIC crime in the last five years’ (CIO). To get to six hundred individuals from these fifteen thousand, the model first applies two variables on police data, selecting those individuals who have been apprehended as a suspect in a HIC event in the last five years and have been suspected of a felony in the last two years. This step reduces the total number of HIC offenders present in police databases to ‘a group of two to three thousand names, which is then tightened up a bit by making the periods a bit smaller or shorter’ (CIO). While the variables are set, analysts will manually intervene to further reduce the number of people included in this initial selection. The second step is to cross-reference this list with public prosecutor data. Thus the the Top 600 selection are those individuals who qualify on the basis of police interaction, have been arraigned before an examining magistrate in the last five years, and have had contact with the public prosecutors at least three times and convicted once in the last five years. The output

[10]GGD is the Dutch public institution on medicine and health services.

is presented to all of the involved public authorities, and after debate and deliberation, a final list is created.

In my mapping chapter, I observed how the crime landscape is changing: HIC is on the decline and violent offences are on the rise, and high-profile incidents demand a response from the police. A dynamic that is also visible in the Top 600 is that the reduction in HIC rates does not close down risk approaches but rather opens them up to include new crime priorities.

> Six months ago, we ended up with a reality where the HIC problem is somewhat reduced, and drugs and violence became much more prominent. That was a very complicated one by the way because the images from the administrators were not very clear at all; they say there is a small group of big-time drug criminals and cartels who recruit relatively young, influenceable men, with minor criminal offences on their name, to sell drugs. Within no time, they are given a gun and are involved in liquidations or told to hang a hand grenade on a wall, a door handle, or something like this. (CIO)

This expansion required the Top 600 to define a new set of criteria that would allow them to identify and select those young men who are most vulnerable to being exploited by drug cartels. The evolution of the Top 600 demonstrates how the political response to the rise of high-profile violent offences is to expand an existing programme to include a new safety problem, drugs and violence. Thus, while the Top 600 emerged from a specific political context, its existence shapes avenues for political action, which in turn changes how data-driven risk models are constructed.

5.1.2 ProKid – the Netherlands

ProKid is another Dutch data-driven risk-scoring model that looks at identifying young adults in police data who are at risk of engaging in a criminal career. This tool materialised from a forensic psychologist's observations and her commitment to the safety of adolescents and young adults; as such, I will first discuss her motivation and then move on to the tool itself. The nature and scope of the police is not static. It is subjected to internal and external efforts to change it. One of these internal efforts is ProKid, which originated from the belief that police should take responsibility for the well-being and future of minors under twelve.

> When I started with the police, that was before 2000, the police did nothing about minors under twelve. They could not be prosecuted and thus they saw it as a meaningless activity. In my field – I'm a forensic psychologist – this is strange because

> people who will be in jail for a very long time all have police contacts before they were twelve, so what do you do, you wait to intervene until a minor is twelve. (forensic psychologist)

Historically defined modes of operation are challenged when new skills and world views become embedded within the organisation. Taking into account that there are a number of criminological insights that inform how early childhood experiences, unequal distribution of resources, and limited life changes breed criminality (DeKeseredy and Dragiewicz, 2018), the forensic psychologist observed how

> there is an intergenerational transmission. Violence is passed on within families but also across families generations. [...] You have certain children who come from certain families, who sometimes also have some limitations themselves or have some little things, where you could say, it's kind of bad that those children first have to start their criminal career before anything actually happens in that area. (forensic psychologist)

Here, she makes the argument that, rather than waiting for a minor to turn twelve, the state and the police have the responsibility to proactively intervene in their lives and positively change their future. ProKid offers a way for the police to step up and take this responsibility.

Since its early days, ProKid has had three distinct iterations: ProKid-12, ProKid Plus, and ProKid-23. I will start with the original ProKid-12 model, which tries to identify children that show concerning behaviour under the age of twelve. While, historically, the Dutch police have not actively engaged with minors, they have data about them, ranging from small incidents and small violent altercations that get reported to the police either by the minors themselves or by others. Analysing this data reveals that 'shortly before major incidents, there are several incidents and there is an accumulation of suddenly a whole number of incidents. You see that there are indeed patterns in that' (forensic psychologist). ProKid-12 analyses the number, frequency, and severity of the incidents and activities[11] of a child under twelve registered in police systems, as a perpetrator, victim, or witness of crime, and the incidents and activities registered to their home address. The domicile variable allows this model to account for the home situation, specifically criminal activities surrounding these minors. It is important to note that, even though ProKid-12 aims to identify those minors who show concerning behaviour on the basis of police data, it is not

[11]Incidents are events in society (for example, a robbery), and activities are actions of the police (for example, neighbourhood controls)

considered a policing tool; rather, it is positioned as a signalling and referral instrument to those care authorities who are responsible for the well-being of minors under twelve in the Netherlands.

To make this tool relevant for policing, a new version was being developed, originally to include young adults under eighteen (ProKid Plus) who demonstrate concerning behaviour, but during the development process, the age range was adjusted upwards to include young adults under twenty-three (ProKid-23). This choice was informed by the introduction of adolescent criminal law in the Netherlands in 2014. This law stipulates that when a young adult between the age of sixteen to twenty-three years is arraigned, a judge can take their maturity and circumstance into account and decide to apply juvenile justice law or adult criminal law (Rijksoverheid, 2021). What seems like a minor change had significant impact on the risk-scoring model. Informed by psychological and criminological insights, who argue that, between the age of twelve and twenty-three, many changes take place in the lives of adolescents and young adults, the developers decided to exclude data about the home address and include data about co-suspects.

> When minors transition from primary school to secondary school, there is a decrease of parental and school control, peers become a key influential factor in their lives, and hormonal changes impact the ability of young adults to think through the long-term consequences of their actions. (forensic psychologist)

The ProKid-23 model no longer focuses on the individuals and the home environment but rather on the individuals and the activities of their peers. More specifically, this model calculates a risk score on the basis of data held by the police on individuals, either as suspects, victims, or witnesses of an offence and on the incidents and activities of their co-suspects. The incidents of co-suspects can only inform the risk score if there is at least one common police registration between them. At the time of interviewing, ProKid-23 was not operational but the evolution of this tool again shows how, once a risk-scoring models exist, there are organisational drives to change and expand its purpose. Where ProKid-12 was primarily seen as a signalling and referral tool to other public authorities, ProKid-23 became a tool that can be used both as a referral tool and as a police instrument.

5.1.3 IOM model – the UK

The IOM model is currently being developed by the West Midlands Police, more specifically by Insight Lab. It is a first attempt to mainstream data analytics capabilities into UK policing. The model itself aims to more accurately identify and predict the probability that a known acquisitive crime offender will move from committing low or middle harm offences to high harm offences

(West Midlands Police, 2019). The term *acquisitive crime* is used to describe offences that derive material gain from the crime, such as theft, fraud, and robbery. In the future, it is believed that the IOM model will support the selection of individuals within the existing IOM approach, which is run by a different department, and combines care and control interventions to move offenders away from re-offending. A practitioner reflected:

> It is being produced from the point of view of essentially allowing offender managers to help out and produce list and saying who it is that they should potentially be looking at next. (data scientist)

It is worth noting that the choice to develop a more sophisticated data-driven risk model for the IOM approach was not driven by a specific need that came from the offender manager team nor from a specific safety problem; it was primarily informed by findings of an external innovation partner. 'If you were to ask what is the most pressing problem for policing that we would set as the first job for our lab to approach, this probably would not be it' (detective superintendent). Accenture, the innovation partner for the West Midlands wholesale digital transformation process at the time, was given access to police data and asked to show which insights could be generated from it. In their analyses of police data, they found a correlation between young adolescents under eleven who were caught committing a group offence and the percentage of those that went on to commit serious violent offences later on in life. The detective superintendent reflects on these findings:

> We thought 'blimey, that is interesting.' We had not really noticed that before. What this model has shown us is that a significant number of those individuals would become involved in serious crime. Now we are not going to do anything with these eleven-year-olds that would be contentious. (detective superintendent)

Accenture's test, as such, showed how data analysis could offer new crime insights that unknown to the police, in other words, that those involved in serious crime have also been involved with small offending in a group at a young age. However, his reflection also suggests that some insights, profiling and early intervening in the lives of minors, are too politically sensitive as a first use case. Thus, the findings of Accenture did not offer the police a target group or crime priority but rather a logic of prediction.

> It demonstrated the power of data to create insights. So we developed that into the offender management model. We decided to look at our existing cohort of offenders to see what the pathways had been for those individuals. (detective superintendent)

The police took this logic and applied it to a less controversial use case, that of identifying those known offenders most at risk of transitioning from low to high harm crimes on the basis of data.

The new IOM model tries to make a more accurate prediction on which offender is most likely to escalate from low to high harm offences by accessing the different police databases and applying ML techniques. A practitioner explained:

> Develop a measure of the harm that they have created through their criminal activity, that then allows you to rank people in terms of how much harm they have created. The second element is essentially about prediction, so we basically create this harm score, and we get, for the sake of argument, low harm and high harm and the model is to predict the likelihood or the probability that someone is going to become high harm, obviously before they reached that level. (data scientist)

To predict the risk that someone will escalate from low harm to high harm in the near future, the team first had to create a baseline of what different categories of harm look like. This was done by correlating fifty variables from the eight police data databases.[12] Data from these different databases allowed the West Midlands Police to correlate the recency, frequency, and severity of the offences to see whether a low harm offender is showing concerning behaviour, like recently being caught committing more criminal offences in a short period of time, that would indicate they are moving to higher harm crimes. Other factors that are taken into account relate to drug and alcohol abuse or if the offender has a known history with drug dealing, firearms possession, and violence. Unlike the Top 600 and ProKid, which emerged from the political, organisational, and individual desire to tackle a specific crime problem, the IOM model is a use case to show the merit of predictive analytics in policing approaches and mainstreaming new skills within the organisation.

5.1.4 Domestic violence machine learning model – the UK

The domestic violence ML model is being developed by the chief superintendent of the Hampshire police and aims to forecast the risk of re-offending in the context of domestic violence. At the time of the interview, this model had no name and was referred to by the chief superintendent as the domestic violence ML model. Risk assessment is a normal procedure in the context of domestic violence. Currently, the UK government, from the police to other agencies and local authorities, use

[12]The databases are crimes (crimes committed), IMS (intelligence), ICIS (custody), PINS (prison notification system), Corvus (intelligence and tasking system), OCG (organized crime group data), OASIS (the event logging system), and DiP (drug intervention programme data). The SAS (stop and search) database was part of the original IOM model but is no longer included in it (Ethics Committee, 2020; West Midlands Police, 2019).

one risk-assessment tool, which is called Dash.[13] There was no specific police priority or crime incidents that led to the development of this ML model; its creation is part of the chief superintendent's PhD research. He observed that one of the challenges of Dash is that its protocols require police officers attending the scene of domestic abuse crime to ask the victim twenty-seven Dash questions. This is not optimal, as it takes place in 'an environment that is not always suitable for an interview, and those officers who use it have a huge variety in their skills and ability to apply the model' (chief superintendent).

Dash national deployment and the varying levels of data quality offer possibilities for improvement and strategic opportunities to develop technological capacity within the police. There are a 'fair few points about that risk-assessment tool that can be improved upon. Hence the reason for going down this application root' (chief superintendent). Prior to developing the model, he started with ten focus groups to include the voices of domestic abuse professionals and foreground concerns they might have about the creation of a new model. This brought to light that, even with all its imperfections, practitioners are in support of Dash.

> It creates a common language around risk assessment and it is a much better way of capturing the voice of the victim, and that it is the victim's own assessment of risk. (chief superintendent)

Rather than replacing something that seems to be imperfect from a data perspective, this process allows practitioners to formulate what is good about Dash. The findings of these focus groups informed the choice to not create something new but layer ML on top of Dash.

ML is seen as a way to more accurately calculate the likelihood someone would commit a domestic abuse offence, and assign a low, medium, or high risk to this individual.

> Low risk, that means that they will not be arrested for domestic abuse offence in the next few years. The next one up is when they are medium risk, which means they will be arrested for a domestic abuse offence, at least one, probably several, over the next few years, but that offending will not be serious. And the final one is high risk, where you will be arrested for serious domestic abuse offences in the next few years. (chief superintendent)

[13]Dash is the Domestic Abuse, Stalking and Honour Based Violence, https://www.dashriskchecklist.co.uk/

High risk indicates the likelihood of committing offences ranging from grievous bodily harm, murder, manslaughter, rape and serious sexual offence, and arson. The ML model builds on Dash by including data from the twenty-seven risk-assessment questions and combines these with police data on the perpetrator, such as 'has the offender previously committed a domestic abuse offence, what was the age of onset of the first offence, how many times have they committed a domestic abuse offence, how many times has the victim been a victim' (chief superintendent). In theory, more accurate risk scores on escalating levels of harm can inform the actions of public authorities that are aimed at decreasing the number and impact of domestic abuse offences.

I will conclude my description of data-driven risk-scoring models in the Netherlands and the UK by summarising my findings. Looking across and between the four different data-driven risk-scoring models, it becomes clear that this function emerges either as the result of a political crisis or in response to an organisational opportunity. The data-driven risk-scoring models are primarily situated as organisational support tools that allow police to identify offenders that fit a predefined offender group or individuals that show concerning behaviour. The examples in the Netherlands and the UK show that this data-driven policing function has three main aims.

- To identify minors and young adolescents who increasingly engage in small offences and as such demonstrate concerning behaviour. The police mandate is to signal concerning behaviour and, where needed, inform other public authorities.
- To identify prolific offenders that are connected to a selected security problem, who are not deterred by the traditional criminal justice system interventions. The aim is to coordinate care and manage crime interventions to structurally intervene in a person life.
- To improve existing risk models to more accurately identify known offenders who demonstrate behaviour that might indicate an escalation from low to high harm for a pre-existing intervention.

Studying data as practice allows me to situate these risk models as the materialisation of top-down approaches to crime that emerge from specific political, organisational, and individual interests. All the data-driven risk models are considered to be a small and initial step within a larger intervention that moves from responding once a crime has occurred to minimising the impact of crime by identifying those individuals who show concerning behaviour in police data and actively intervene in their lives through a care and control approach. The origin stories further foreground how risk scores are developed from the assumption that something is not working as well as it should in contemporary policing, the failure of traditional deterrence tactics to change the behaviour of a

small group of offenders, the lack of police involvement in the lives of minors under twelve years old, the idea that the 'right' people are not selected for a specific intervention, or the lack of data capabilities within the police. This suggests that, while the concept of risk is predominantly understood in relation to managing and controlling the actions of a flawed individual who is inclined to commit a criminal offence in the near future, my findings suggest that the police see it in part as the result of their inability to prevent it. In the next section, I will further explore this argument by engaging with the multitude of meanings inscribed into risk.

5.2 Risk as an internal organisational optimisation logic

In the scholarly debates discussed in chapter 2, predictive identification is primarily situated within the managerial logic of datafication, where the process of attributing risk to an individual should allow police to pre-emptively intervene in their lives and thus become more efficient and effective. This is often referred to as police being able to do more with fewer resources. This, I will argue, offers a flat ontological view of the function of risk. Thus far, the origin stories of the four risk models foreground how crime is seen as the result of the individual committing it and the inability of the police to prevent it. In this next section, I will continue to explore this observation by drawing on my empirical data to argue that practitioners inscribe different meanings to the construct of risk: a normative measure of behaviour, a normative label that justifies a specific intervention, and a normative construct that allows for a range of state institutions to coordinate their activities. Although I engage with them as isolated constructs to highlight their specific affordances, they should not be seen in isolation from each other but rather as a multitude of meanings that are inscribed alongside and in relation to each other. I will conclude this section by arguing that these different notions of risk reveal that police primarily turn to risk scoring to optimise for specific organisational needs. Thus, it demands a more nuanced understanding of to what end these technologies are introduced, beyond the managerial logic of increased efficiency and effectiveness, and towards what I refer to as the internal *organisational optimisation logic*.

5.2.1 A normative measure of risky behaviour

I will start the exploration into what meaning is inscribed to the construct of risk by engaging with the question of what knowledge gets privileged in data-driven risk-scoring models. In chapter 2, media scholars theorised how the datafication of society created an ontological shift on what is considered authoritative knowledge, one that privileges correlation over causation. My findings resonate with this observation, and I will build on it to argue that the political rationale that is

enacted through these risk models offer a one-dimensional view of someone's life, where their negative behaviour is magnified, and the positive elements in a person life are undervalued.

A practitioner observes that risk scoring is a way to move from a descriptive to a prescriptive problem statement, in other words, moving from an abstract problem to practical variables. Take, for example, the recent changes to Top 600, where the programme was asked to expand their focus by including a new safety problem, drugs and violence. As previously explained, local politicians have expressed the desire to identify and change the life course of those young men who need to be made more resilient to the temptations of organised crime. This political ambition in itself creates a normative understanding of a problem, vulnerable young men who need to be protected by the state against immoral drug cartels. However, in the context of this thesis, I will engage with this politically defined problem by pointing to the ambiguities that become embedded within a risk score when one tries to operationalise an abstract problem.

> This is not an exact science. These are politically defined problems: young men who are vulnerable to the influence of drug criminals, who might be sent out with a weapon. These are very soft descriptions that you try to operationalise [into hard criteria]. (CIO)

To operationalise the drug and violence priority, a risk profile is created along the lines of all individuals under thirty that are in the police registration system and have been arrested in the past five years for selling hard drugs and committing a violent crime, threat, or being in possession of a weapon. This approach places a normative measure of behaviour on the selected individuals, as, in the eyes of the state, they are no longer young men who have a range of challenge and opportunities but are labelled as young men who are at risk of being recruited by drug cartels.

In the British context, both ML models aim to more accurately predict the escalation of low harm to high harm offences. The underlying assumption is that, once these models are completed, they will add value to the existing programmes by improving the selection process. When discussing the IOM model, the detective superintendent reflects that,

> as soon as you start talking about predictive policing certain people start referencing things like Minority Report and they think that it is outrageous, while actually, we have been managing offenders based on a risk score matrix for quite a long time, and all we are trying to do here is make it more sophisticated and accurate. (detective superintendent)

Risk scoring, as such, is not a new practice, but rather, data allows for the optimisation of existing practices. Where the concept of risk and the ability to better and more accurately calculate the riskiness of someone escalating in levels of violence on police data is a given, a neutral fact.

> The accuracy of our model, it gives a very very strong signal, so it is a very strong model. It is weighted towards specificity rather than sensitivity, so because we are looking for people, trying to correctly identify those people who are not going to transition. (lead data scientist)

This requires an articulation of what knowledge gets privileged when we view society through a police understanding of risk. All the models build on what is already known to the organisation, more specifically machine-readable data points of incidents and activities in which individuals have come into contact with the police and or public prosecutor. Several practitioners observe that the information position of the police is unique, as it is the only public authority that has the mandate to enter into the private sphere of all layers of society when there is a suspicion of criminal activity.

> The police in the Netherlands, as the public authority operational on the ground across all layers of society, holds a unique information position that can best inform these decisions. (district chief)

> That we are also one of the few companies that look behind people's front doors twenty-four hours a day. (district chief)

However, risk that is determined on this information position, as practitioners themselves observe, will offer a one-dimensional view of someone's life. It will drill down on negative behaviour that is observed and registered in police databases and excludes positive elements that might be present in someone's life, which can range from a conducive family environment, a job, education, 'good' friends, or the presence of a trusted authoritative figure.

The risk gaze that emerges from police data as such will approach a person as a potential criminal and not what this person is able to do or be within a certain context. Some practitioners argue that the models are only an initial step in a larger approach and that these protective elements will be identified when the human steps into the decision-making process. Others argue that protective elements are attributes that are implicitly built into the models; their influence will prevent an individual from coming into contact with the police, and as such, they will not get selected for or

naturally be removed from the list. The last assumption is contested, also among practitioners, as it is known that having more resources and interventions directed at an individual will, by default, make them more visible to the police. In defining risk as a more accurate measure of potential criminal behaviour, we run the risk of entrenching a security view onto how these individuals are perceived and treated. Therefore, I argue that the concept of risk in the context of policing creates a normative measure of risk behaviour. Even when the selection is made for a combined control with care, changing the live conditions of a risky individual, individuals are selected on police data for the purpose of crime prevention.

5.2.2 A normative label to justify a specific intervention

In this section, I will explore what materialises when normative measures of risky behaviour become embedded within a policing approach. I will structure these insights along my conversation around the Top 600, as it is the only risk model that is currently in use, and argue that the variables not only select individuals for an approach but also form the justification regime for the entire intervention. I will start from the observation that those working in the context of the Top 600 problematise the idea that risk models allow for accurate risk calculation. The creation of a risk model is dependent on the process of translating politically defined descriptive problems to operation criteria. This process is informed by the relevant police and criminology literature, historic data on the young men who fit this description, and in consultation with the relevant public authorities. The CIO explains the process on the Top 600's original HIC criteria:

> Someone who is on the Top 600 has a history of criminal behaviour whereby the chance of re-offending is there, but it is not always certain how big this risk exactly is. (CIO)

The CIO's observation suggests that data-driven risk scoring as such is primarily used to identify known offenders who show certain behaviour rather than actually determining the risk of re-offending. When applying a data-driven model to an operational situation, specific challenges emerge.

> So there are guys you had the feeling that we should actually take them along. And there are guys who you felt maybe we should have included a little less – it's not like that's a hard false positive or false negative. (CIO)

These models are able to identify known HIC offenders, but they are less capable of predicting the actual riskiness of the specific individual. Scholars and technologists refer to this as *false positive*,

unjustly being included in the output, and *false negative*, unjustly being excluded from the output, but the CIO argues that reality is more complex than these binary understandings of data modelling.

> There is a soft target group just below the Top 600, the slightly less prolific offenders, who have just as many problems as the identified individuals. So if there's a slightly less active HIC multiple offenders in there, and a slightly more active HIC multiple offenders falls out, you wouldn't have wanted to swap them. Because that slightly less active multiple offender who's approaching [the list] can still be bursting with problems. (CIO)

Arguably, this 'soft target group' still fits the criteria of the Top 600 and, when ranking between eight hundred and one thousand individuals, there might be those who are more prolific and should theoretically qualify for the Top 600 over others, but all of them have problems. All who are in this 'soft target group' are believed to benefit from the care and control interventions to manage and reduce their criminal activity and to slightly improve their circumstances. These insights challenge the notion of false positives, as the boundary between who falls within and outside the Top 600 is not hard but diffuse. There are a few false negatives, where the team knows about individuals who should qualify for the approach, but 'who in the past have been charged with the wrong article, a non-HIC offence such as kidnapping or a drug rip deal' (CIO), and as such do not meet the criteria. However, as the CIO explains, this is not enough reason to make exceptions.

> The moment you start making exceptions, the criteria no longer apply, and that really is the basis for your approach. That is your legal justification for putting people on that list, so you can't deviate from that. (CIO)

He articulated that risk and the variables on which it is computed are an organisational process to externally justify why certain people qualify for a specific intervention. This indicates that variables in a model have a two purposes: to select individuals for a predetermined approach and to justify the selection and approach.

5.2.3 A construct to coordinate between state institutions

So far, I have shown how practitioners inscribe meaning to the construct of risk. It allows them to identify risky behaviour and justify an approach. I will now explore the third meaning that practitioners ascribe to risk, that of coordinating action across and between different public institutions. Risk as a construct for coordination stems from the belief that the police are not

necessarily the best actor to act upon the risk score; for example, the police have the mandate to act on a criminal offence, but care and child protective authorities might be in a better position to change the conditions surrounding a child, and the other way around. Unlike the care and child protective authorities, the police has the authority to show up unannounced and enter an individual's house. The belief that underpins these coordination efforts is that, by investing in collaborative efforts around a specific security problem, the state can better intervene in a person's life. Practitioners observe how individuals who end up getting selected for a specific intervention through a risk-scoring model often have complex problems and a long, fragmented history with a range of state institutions.

> We do know that the problems in a person's life in many areas are so broad and encompass so many domains that a government needs to know that people who are concerned in one domain are also concerned in another domain. (policy adviser)

Similarly, another practitioner said:

> If someone has problems in their family or is struggling with substance abuse, health problems, or has debts, and also re-offends, then all these elements should be reasons for public authorities to coordinate their work to improve the conditions in someone's life. (CIO)

Therefore, practitioners believe that a sole focus on the 'riskiness' of the individual flattens what they are trying to achieve and does not acknowledge the fact that these individuals are often already engaging with and dependent on a wide range of public institutions. Here, risk scoring is positioned as a construct around which public authorities can coordinate their efforts, step up, and take more responsibility.

Risk, as such, might be calculated on an individual, but it is a construct that allows public authorities to coordinate their approach under one umbrella, which is referred to in the Netherlands as a 'one government' approach and in the UK as multi-agency partnerships. What characterises these approaches is that local authorities, police, care, education, and welfare authorities regularly meet and jointly intervene in the lives of individuals.

> These authorities discuss who is in the best position, [and] who holds the best resources and mandate to intervene. (district chief)

Data-driven risk-scoring models as such can be understood as a social construct that informs the initial selection, after which a range of public authorities join forces and invest resources to intervene in the lives of the individual and their family for a longer period of time. When an individual is identified as 'risky', the idea of the 'one government' approach is to make one authority responsible for the implementation and coordination of the interventions of all. Unlike the popular belief that predictive identification will allow police to do more with fewer resources, a practitioner observed that these integrated interventions 'costs a lot of time and energy' (CIO) and are very resource-intensive.

I will end with reflecting on the Dutch police pyramid of increasing complexity, which is a prime example of how risk as a social construct allows public institutions to coordinate interventions and jointly respond to a security problem. A police practitioner explained in a police meeting on ProKid-23 that the pyramid has four layers: the bottom two layers consist of, give or take, 80 per cent of the population, who are doing fairly well and have little to no interaction with the police. The role of the police in the second layer is to be observant, keep an eye out for those individuals who run the risk of escalating to more complex problems, and, when needed, refer them to relevant care authorities. The third layer consist of individuals who increasingly engage in small offences and as such have more interaction with the police; the focus is on safety and care, and interventions are aimed at positively changing the conditions of people's life.

> The top 5 per cent of the pyramid is occupied by individuals who are known throughout the criminal justice system and create insecure situations that require immediate emergency police responses, Police efforts are primarily aimed at limiting the negative consequences of crime through enforcement; the care and safety mandate is of secondary concern. (adviser)

She continues to explain that, in the context of domestic violence, situations quickly escalate to the fourth level, and risk assessment supports the decision-making process on whether to intervene with immediate police action or whether the situation can still be managed by the responsible care authority. In contrast, minors are often lower in the pyramid; here, instruments like ProKid allow the police to identify, signal, and respond to individuals who show concerning behaviour. Finally, there are the Top X approaches, as in Amsterdam the Top 600, which aim to identify the most prolific offenders who populate the third and fourth layer of the pyramid, and interventions are directed at minimising crime through a combination of managing crime and care interventions. A

common critique of ‘predictive’ policing programmes is that, even when the intentions are public health-oriented, they end up being punitive. What the pyramid of increasing complexity and the practice of risk modelling in the Netherlands shows is that punitive actions are considered to be part of the intervention, just like signalling, care, and coordination efforts. Which actions are taken depends on the individual, their situation, and the level of harm of the offences. The pyramid of increasing complexity again points to the centrality of risk in coordinating action across and between different public institutions.

I will conclude this section on the meanings practitioners ascribe to risk by summarising my key findings. Contextualising data-driven risk scoring allows me to situate this development within their broader environment and move us away from the flat ontological view, in which risk is merely tied to an individual, towards one in which the construct of risk is multifaceted and used as a normative construct that allows police to optimise and justify certain practices. In this section, I discussed three notions of risk that emerged in my empirical research: a normative measure of behaviour, a normative label that justifies a specific intervention, and a construct that allows for a range of state institutions to coordinate their activities. All three point to the dynamic that data-driven risk scoring emerges from the belief that the origins of crime are the result of the flaws of the individuals and the failures of the police and the broader state to manage these individuals; further, the belief is in the promise that data analysis will allow police to transform and overcome their internal challenges. This requires us to see risk as a normative construct that primarily aims to optimise certain organisational processes, a dynamic that I will refer to as an *internal organisational optimising logic*.

5.3 Risk as a driver of police practice

This chapter has thus far explored the use of four risk models and foregrounded how police practitioners inscribe multiple meanings onto the construct of risk, which creates a dynamic of internal organisational optimisation logic. These insights contributes to my research interest to explain the actual nature of data-driven policing and how it is changing the ways practitioners understand crime. In this next section, I will argue that the sociotechnical imaginary (Jasanoff and Kim, 2015) of risk and data places normative expectations, both internal and external, on the police. This gives rise to a number of policing practices, such as a responsibility to intervene, innovation, and safeguarding. I will conclude that these normative expectations invoke a sense of inevitability, where for practitioners, it is not a question of if but under which conditions data-driven risk scoring can become embedded within the organisation.

5.3.1 A responsibility to intervene

Research into the use of predictive location policing has suggested that these technologies allow the police to shift from a more reactionary approach, acting once a crime has occurred, to a more pre-emptive policing approach, intervening in someone's life before they have committed a crime (Brayne et al., 2015; Van Brakel, 2016). In this section, I will draw on my research to contextualise this changing nature of the police to argue that, from an individual police officer's perspective or within a specific intervention, this function is seen to open up avenues for action, but from an organisational standpoint, its use is contested, as it runs the risk of, at times, placing unattainable or unwanted normative expectations on the police to pre-emptively intervene.

Some police practitioners feel that they have a duty of care to prevent certain criminal futures from unfolding. In this view, police can take on the role of the authoritative figure in the life of those minors and young adults that show risky behaviour, which should deter them from continuing into criminal activity.

> You could also think we are the police, and we are also interested in investigatory possibilities or that you say, to a certain extent, children who become criminals lack supervision and someone who, on the basis of authority, tells them what it means if you commit a crime. Don't we have a job in representing authority, which sometimes they lack even at home? (forensic psychologist)

The assumption that underpins this shifting role is that the state is the legitimate authority that is responsible for creating equal opportunities in society. Data-driven risk scoring supports this mandate by better identifying those young adults who are in need of additional attention. Those individuals who are situated within the youth, care, and security policy area of the police, however, argue that the preventative task is nothing new: 'The police youth task has traditionally been prevention and early detection first and repression last' (policy adviser). It is the turn to data-driven risk scoring that is seen to have given rise to debates on the role of the police in care and control intervention, specifically in relation to the nature and scope of other public institutions.

> Prevention is already deeply embedded in the youth task, and what you see now is that the discussion arises – you see it is now shifting in society. What is our task, how far should we signal, and how far is this now actually a task of chain partners, municipality, schools, when you talk about youth? (policy adviser)

These concerns are echoed in discussion with other police practitioners,

> At the time ProKid-12 was developed, the police had no issues with it. Now that data-driven risk scoring is getting a different form – the age range is extended to twenty-three – this raised several concerns. Do we want this? Is this a police task? Can we justify that we do this? And if we identify that something could happen, what does this mean, will someone do something with this signal? (strategic adviser)

Thus, the introduction of technology is seen to reopen questions around the role of police to pre-emptively intervene.

Practitioners feel that the introduction of data-driven policing creates the unrealistic expectation that, through prediction, the police can prevent harms from materialising. As one practitioner argued, the idea that if a data-driven risk score signals that someone is showing concerning behaviour, that does not necessarily translate into action. ‘There’s a whole lot about keeping a grip, but the discussion about how realistic is it, that’s not being had’ (strategic adviser). Again, drawing the conversation back to an organisational challenge, where police are confronted with the societal desire to control and eliminate risk. Other practitioners observe: ‘after all, we live in a society in which we have had some really serious incidents in which the politicians tend to look at who is responsible’ (district chief). She points to an increased desire and expectation of control in society, where, in the past, the political and media response to high-profile violence incidents has been to quickly shift the blame from the immoral perpetrator to the public institution that ‘failed’ to prevent the incident from unfolding. This is a culture that has instilled fear in the senior ranks of the police, fear that their institution will be held responsible for not properly intervening in the lives of those individuals who are showing risky behaviour. Foregrounding a dynamic in which, even though police are actively investing in the development of these tools, as part of specific interventions or within certain departments, there is a managerial hesitance that (temporarily) constrains its use.

Another point of friction lies in the just-do-it policing culture; middle managers are reluctant to introduce this function as it will create the presumption to intervene (Fussey and Murray, 2019). Several practitioners note that it is in the nature of the police to respond to a situation. If a risk score indicates that certain minors and young adults show concerning behaviour and no one steps in, police officers will be inclined to try and intervene.

> The nature of police is to take a step forwards when everyone else takes a step back. (district chief)

> We once had youth and community police officers. At the end of the eighties and the beginning of the nineties, we went too far in that we actually became more of a social worker than a policeman. Now the police are more concerned with enforcement and detection than with providing assistance to families. You notice that the police chiefs fear that, if they deploy risk tools, then their neighbourhood police officers will certainly go all out to help families again when they feel that people safety is at risk. (district chief)

Middle managers fear the introduction of data-driven risk scoring will encourage police officers on the streets to take on too many care-related responsibilities. This is concerning, as police do not have the mandate or expertise to intervene in a range of problems.

> We're not going to diagnose or assess, that requires different expertise. The situations young people end up in is what triggers you to make a care report. (policy adviser)

This points to a tension between the normative expectation to respond – invoked by data-driven risk scoring and the observation that the police can observe, document, and report – but should leave the diagnoses of mental problems and care interventions to the professionals who are trained for it.

What we can learn from these reflections is that the emergence of data-driven risk scoring creates tensions within the police about their role in society. Individual police practitioners who develop these data-driven risk-scoring models have a strong sense of a duty of care to prevent certain criminal futures from unfolding. Yet, strategically, data-driven risk scoring creates the internal and external expectation to intervene, which, from an organisational standpoint, risks placing unattainable or unwanted normative expectations on the police. Thus, the introduction of this function within policing can be perceived to close avenues for not intervening, as it creates perverse incentives to act if merely to avoid being held responsible for when a risk materialises. Operationally, the introduction of these risk-scoring models is feared to legitimise police officers on the street to 'step up' when no one else does and intervene in an individual's life regardless of whether they are trained to deal with the complexity of the situation.

5.3.2 Risk as innovation

The second practice I want to draw attention to is risk as innovation. In my interviews, I found that practitioners see the practice of data-driven risk scoring as an opportunity to push through a broader organisational transformation, upgrade their data infrastructures, and include new skills within the organisation. In this section, I will position the turn to data-driven risk scoring as a practitioner response to the normative expectations that the datafication of society places upon the police. A professional organisation in the twenty-first century is data-driven, which requires the police to optimise for innovation, update their data infrastructures, and invest in a new skill set.

A general observation in my interaction with the police is that practitioners are not immune to the promise of data, where the datafication of society creates normative expectations about what a professional police force should look like, and the use of data systems are believed to make their organisation more efficient and effective. Where the practice of data-driven risk scoring, at first sight, speaks to the latter, doing more with less, my interviews revealed that individual tools are primarily seen as a vehicle to create organisational momentum and resources to adhere to the first. A practitioner described the innovation process of the West Midlands Police:

> Accenture won that contract, so when we started developing Data-Driven Insight, we developed a proof of concept, and we used their data scientist and engineers. And once we demonstrated that it was feasible and had an added benefit, then we went to the recruitment process to make that sustainable, ‘cause it is very costly when you outsource that all the time, and you risk losing control as well. (detective superintendent)

External expertise is brought in to test an assumption – such as whether data analysis will allow police to gain new insights and whether this will have an added benefit for the operation of policing – which may justify an investment into constructing the in-house data analysis team. As previously mentioned, Insight Lab is expected to mainstream data analytics capabilities into policing and make its practice more cost-efficient, and the IOM model is their first project. ProKid and the domestic violence ML model also use external actors, primarily universities, to create their models.

> One of my learning points from this is that, for anyone who commits to doing this sort of work, has to do that with academia. I think it has to do that with commercial enterprise, both from a business consultancy perspective but also from a technological

> capability as well. Law enforcement hasn't got the tools to do that on its own, not yet anyway, so the only way you can do that is to collaborate. (chief superintendent)

Collaborating with external partners thus is seen as a practice that will allow the police to borrow skill sets from other sectors to create a proof of concept and make the case for the allocation of additional resources to a specific tool or skill set.

The desire to innovate has consequences for the operational side of policing. Practitioners observe that the organisation itself needs to be optimised for the twenty-first century by investing in their critical data infrastructures and in-house skills. The lead data scientist who was recruited for Insight Lab and moved from the commercial world into policing shared his observations on the differences in the work environments:

> The technical IT environment and processes are not necessarily designed in an optimal fashion from a data science point of view, so obviously we are working through a few technical issues. (data scientist)

Here, the West Midlands Police first had to invest in the interoperability of their database before they could start to develop the ML tool. Innovation as such is hampered by the infrastructural reality of policing, where the historically grown data infrastructures are not necessarily suitable for more advanced data-driven policing tools. Another practitioner highlights a similar tension between the possibilities of technology and the reality of police data infrastructures:

> I think that the police's own baseline of technological capability is not great. We spend a lot of investment and a lot of time actually just trying to keep the lights on, supporting sometimes old and antiquated systems or systems that do not talk to each other. So that is where a lot of our investments have to go to. Investment in future technology is not as great as it can be. (chief superintendent)

The emphasis on old and antiquated data systems shows how building data-driven risk-scoring models is more than just applying ML techniques in a test environment; it requires police to first invest in the quality and interoperability of data and their data infrastructures.

I will end this section on innovation as a practice by reflecting on my findings that building and testing data-driven risk-scoring models in themselves are not a guarantee for future investment in

and deployment of these tools within the organisation. Police have limited budgets and competing demands on them.

> I will find out in the next couple of weeks if my police force, among all our other investments we need to make in digital, has got the capacity to do this one as well. (chief superintendent)

Another practitioner observed how the awareness of competing budgetary demands and the cost-benefit considerations when investing and embedding innovation within the organisation is dependent on the positionality of the practitioner. Data-driven risk models that are created in a silo by 'techies and idealists' do not necessarily understand how to sell innovation within the organisation.

> You are so busy trying to scientifically prove that you're right that you forget there are a lot of people in the police organisation who aren't interested in that at all. Take, for instance, domestic violence: if you can't [or] don't demonstrate what it costs society, then nobody is going to want your risk-scoring tool. It's not interesting at all that you're developing it, unless you show it's societal cost and why you're developing it. (district chief)

This highlights that, when the development of data-driven policing is confined to police practitioners and third parties who are detached from the operational side of the police, they fail to see that the introduction of any new practice is often a simple cost-benefit analysis. She continues to argue that

> we live in a society where we look at what it actually costs and decide if it is worth it. If you get them to see the usefulness and necessity outside your ideological world, then people will go along with you. That's what we sometimes lack: people who can make the bridge between science and what you can actually do with it in practice. (district chief)

The insights in this section reveal that there is a clear desire on behalf of individual police practitioners to innovate. In the case of the IOM model, it is also a use case to show the added value of building data science capacity with the police force. Innovation as a practice builds on the belief that a professional organisation can use data and its affordances to the organisation's advantage. However, this practice runs against a number of organisational challenges, which range from

outdated and incompatible data infrastructures to the struggle to show its added value to the operation of policing.

5.3.3 Safeguarding as an emerging practice

So far, I have explored the practices of intervention and innovation that emerged through the introduction of data-driven risk scoring. In this section, I explore the third practice: that of safeguarding. I will start this section by engaging with the different safeguarding practices to explain what police feel is at stake when data-driven risk scoring is introduced within an organisation. Here, I observed that its emergence is seen as inevitable, and practitioners feel that it is a question of finding the right conditions under which these tools can be introduced so that their use won't negatively reflect upon the organisation. This has given rise to a number of safeguarding practices that outsource the solution to possible conflict to the technology, enacting the right thresholds and external oversight. I will conclude that these practice emerges from the assumption that these tools will optimise for existing practice, and as such, the police merely need to account for the negative externalities that materialise from the use of data, not from police actions.

I will start the section on the actual practices of safeguarding with a closing remark of a Dutch police adviser, in response to when I asked her if I missed anything in our interview. In her response, she said:

> Can we still do without the risk-assessment tools; can we not to use it? That question I think you just have to answer no. So we have to. It's possible, so much can be done, so what we have before us is not the question of whether or not to use ProKid-23, but how are we going to use it in a responsible way? How do we take society along with us so that they don't lose confidence in the police and other government agencies? (policy adviser)

The mere existence of a certain technological function, as such, moves the discussion away from whether to how the police can ethically use it so as not to lose public trust and confidence. In this light, I want to discuss three safeguards that emerged from my interviews. The first is testing for non-statistical bias to uncover how police data is distributed and whether it affects ethnic minorities differently than other parts of the population.

> There is obviously a concern that there may be biases in the data against various ethnicities as a result of previous biases on behalf of police. (UK data scientist)

This particular safeguarding practice aims to account for the fact that certain communities are over-policed and over-represented in police data and outsources the idea of fair policing to technology. There are a number of limitations to this approach: first, it is important to acknowledge that this type of safeguarding is only possible for police forces that collect ethnic data as part of their reporting practice, which is not standard practice for most European police forces. Second, while this safeguard allows police to identify possible bias within police data, it does not account for the fact that the police gaze embedded within data-driven risk scoring is socio-economically determined. For example, all these approaches privilege HIC or acquisitive crime over, for example, white-collar crime. The solution as such runs the risk of further obscuring the normative understanding of crime, as it moves the question away from what is fair policing to what is fairness within a technical artefact. The adviser to the PCC explains:

> I worry that if we have something that is simply more efficient, then it might end up just speeding up all kinds of punitive actions rather than preventative, rehabilitative or public health responses to crime, which is just inherently bad for all of the reasons we know in the criminal justice system is flawed already. (oversight)

He challenges the premise of these tools, that more accurate predicting will lead to better policing, by arguing that policing is not fair to begin with, and increased interaction with police and the larger criminal justice system is not necessarily desirable, wanted, or believed to have a positive impact on the lives of individuals.

Another safeguard foregrounded by practitioners is adding specific thresholds in the data model to justify the approach. The Top 600 uses a relatively simple model that is based both on police and public prosecutor data. The initial Top 600 list was created solely on police data, but the public prosecutor criteria were added not long after. The CIO explains that ‘the throughput times of the public prosecutors’ office are long’, and once an individual gets convicted, they are well on their way in their criminal career; as such, this criteria allows the model to select those people who are responsible for the majority of HIC offences. Here, it is acknowledged that any approach that aims to identify an individual as part of a security problem should ensure that the risk scores are driven by the behaviour of the individuals and not by the behaviour of the police. In this sense, the public prosecutor data provides another threshold to ensure that people are identified on the basis of their actions and not on the basis of the police. It is important to note that the public prosecutor variables are seen to ‘give the selection of the individuals more weight and a certain legitimacy’ (CIO). As

such, this safeguarding practice has a two purposes: preventing the inclusion of people on the basis of police behaviour and giving the output more legitimacy.

A final safeguarding practice is that of outsourcing checks and balances to an independent oversight body. The West Midlands PCC office established an independent ethics committee to provide advice to the PCC and chief constable on the ethical challenges of data science projects such as the IOM model. As outlined in chapter 4, the UK introduced regional oversight in 2012 through elected PCCs, who are responsible for the budgets and oversight of police activities, which includes signing off on data-driven policing technologies funded through local budgets. For the West Midland PCC, the ethics committee is 'more than ticking a box': they have brought together a panel of experts, composed of human rights, legal, data protection, public sector, and police experts as well as community representation, that hold sufficient expertise to evaluate the merits and harms of particular technologies throughout their development and deployment process. The IOM model is the first project Insight Lab submitted to the ethics committee. All the associated briefing notes, agenda, and meeting minutes are publicly accessible on the PCC website. Initially, Insight Lab placed an emphasis on the quality of data science and the accuracy of the model, and the committee came back with concerns related to the freshness of data and the integration of specific databases in which communities of colour are over-represented, like the SAS database. There were also questions about how the tool would be integrated within police operations. The ethics committee is consulted at different stages of the development process, and all information about the models and the responses of the committee is publicly accessible; as such, the mitigation is outsourced to a group of experts and the public at large.

To conclude this section, I will argue that the organisational optimisation logic is a response to the normative expectations datafication places on the police and creates the conditions for a number of practices to emerge. Specifically, intervention, innovation, and safeguarding are all practices that are seen to open and close policing futures. While these practices speak to a changing role of the police in society, which has primarily been discussed as moving from reactive to pre-emptive policing (Brayne et al., 2015; Ferguson, 2019, 2017; Van Brakel, 2016), I argue that this generalisation does not fully capture the nuances of what it means for police to become more pre-emptive nor the internal conflict about whether the police should take on this role in society. Embedding data-driven risk scoring within the organisations is believed to create a dynamic in which police feel they need to act upon its output, invest in the underlying data infrastructures and technological expertise, and establish certain checks and balances. For, the practitioners it is not a question of whether they should engage with data-driven policing practices but rather under which

conditions these tools can be introduced so that their use won't negatively reflect upon the organisation. These practices are based on the assumption that risk scoring will allow police to optimise for existing practice, and as such, they merely need to account for the negative externalities that materialise from data – not from policing itself. Therefore, I will conclude that, while data-driven risk scoring is often positioned as a sociotechnical system that can attribute levels of risk to an individual, it should primarily be understood as a construct that is shaped by and is shaping policing as an organisation.

5.4 Conclusion

In this chapter, I explored four different data-driven risk-scoring practices in the Netherlands and the UK to create knowledge about the actual nature of data-driven policing and its impact on how police practitioners come to understand crime, police power, and justice. It becomes clear that police are actively testing and experimenting with this function. Looking across and between these different implementations allows me to distil that, while three out of four models are still in the development stage and might never become part of the operational side of police, the logic of data-driven risk scoring and the promise of pre-emption is becoming embedded within the organisation. When approaching these models in isolation, we can see that they are a small and initial part of larger crime intervention approaches, in which the concept of risk externalises the origins of crime to a flaw of the individual who commits it and of the police who do not prevent it. When we place these same models within the broader context of policing, other dynamics emerge; this allows me to move away from the flat ontological view of risk that is merely tied to an individual towards one in which the construct of risk is multifaceted and primarily used as a normative construct that allows police to operationalise and justify certain practices. Risk is used to create a normative measure of behaviour, a normative label that justifies a specific intervention, and a construct that allows for a range of state institutions to coordinate their activities. These insights inform my thesis that risk can be understood as a normative construct on which the rights of individuals, the safety and security mandate of the state, and organisational and political priorities are continuously negotiated. I refer to this as the *internal organisational optimisation logic*. These findings have broader implications for how police power and justice is understood in a datafied society, which I will theorise in chapter 8. I conclude this chapter with the observation that there are three distinct practices that emerge with the introduction of data-driven risk scoring: intervention, innovation, and safeguarding. These practices emerge as a response to the normative expectations that are placed on police with the emergence of data-driven policing. Thus, datafication more generally and the introduction of specific tools create a dynamic that can both legitimise and challenge the police as an authority, and

practices emerge to pre-emptively mitigate the negative externalities that arise from the use of data-driven policing.

6. The practice of recognition technologies

The use of biometric recognition technologies by police has been subjected to significant public critique; for example, the campaigns, particularly in the United States, UK, and Europe, that have called for either a moratorium on their use or an outright ban on their presence in public spaces (Big Brother Watch, 2019; Kind, 2019; Reclaim Your Face, 2021). In this context, Brayne and Christin (2021) argue that debates on the use of data-driven policing are overwhelmingly speculative, and at times very little is known about its actual practice. Even when the public knows about and critiques, the use of a specific facial-recognition system, it is unclear how these tools are actually being used and how they come to shape the way the police understand and act on crime. To contribute to my research questions on the actual nature of data-driven policing and the relationship between datafication and police power, this chapter will present my second case study, on the practice of biometric recognition.

Biometric programme	**Police**
Catch, a facial matching technology used for forensic investigation	Dutch national police
Digitale Perimeter, pilot with on-the-edge facial recognition	Amsterdam police
Pilot with facial recognition at Zaventum airport	Belgium federal police
Preparing for biometric recognition	Danish federal police
SIIP, research project on the use of voice identification	Interpol, Metropolitan police (UK), Ministero Della Difesa (Italy), Ministério da Justiça (Portugal), and Bundeskriminalamt (Germany)

Figure 5: Overview of biometric implementations discussed in this chapter

In this chapter, I will present the findings of my second case study on the use of facial recognition and voice identification in Europe by placing the function of biometric recognition under scrutiny. My findings are based on twelve semi-structured interviews with police practitioners in the Netherlands, the UK, Belgium, Denmark, and Europe, as well as participant observation in police meetings and the study of grey literature. The Danish practitioner is included in my sample, as he is vocal in European policy debates. The five biometric recognition systems (see details in figure 5) were selected, as they are stand-alone projects and are introduced with the aim to bring something new to how the police operate. The term *biometric recognition* is used to make a distinction

between more traditional forms of biometric investigation, such as fingerprinting and DNA analysis, where evidence collection is restricted to a physical crime scene at a specific time, and newer forms, such as facial recognition and voice identification, that allow police to capture a digital representation of a body part after a crime has occurred or in real time from a wide range of public and private data infrastructures for forensics, investigation, and intelligence.

This pan-European research does not intend to offer a comparative analysis between the distinct deployments, as it will lack the breadth and depth needed; instead, it aims to identify broad organisational principles that structure police approaches to technology to shed light on the social structures that (re)produce them. This chapter is organised as follows. First, I present the emergence of five different face comparison and voice identification use cases. Second, I will explore to what end biometric recognition is constructed. My research finds that practitioners inscribe multiple meanings to biometric recognition: to automate and attribute meaning to an identity, to justify the expanding nature and scope of the police, and to affirm police as a competent authority. My findings show that this function should primarily be understood as a construct that is shaping how the police engage with their external environment, a dynamic that I will refer to as the *external organisational optimising logic*. Towards the end, I will engage with the practices that emerge from digital biometrics.

6.1. Biometric recognition deployments

This chapter will first explore the practice of using biometric recognition in five different use cases across Europe: Catch, Digitale Perimeter, a pilot at Zaventum airport, preparing for biometric recognition in Denmark, and the SIIP. In contrast to my risk-scoring case study, where practitioners situated its emergence in relation to the origin stories and the actual uses, biometric recognition is discussed in terms of what is happening now and the possibilities for the future. As such, this section will describe current practices, how they is embedded within the organisation, and the possible biometric futures practitioners alluded to. To conclude, while these cases are at different development and deployment stages, all are built on the belief that this function can address certain policing needs, such as identifying an unknown perpetrator by using the increased volume of data that are accessible to police, and possibly in the future identify and monitor individuals in real time for the purpose of investigation and control.

6.1.1 Catch – the Netherlands

In the context of forensic investigation, the Dutch national police are deploying the face comparison system Catch. After an initial pilot, the software was embedded within the police in 2016. The lead biometrics specialist explains how ‘our specialism is forensic investigations, usually from a scene of a crime, so small fragments that one leaves behind accidentally. Here we saw facial recognition, face comparison as we call it, as an opportunity, it is nothing more than a large search engine to efficiently look through millions of face images’. The forensic practitioner positions the use of Catch as a natural progression from their work with fingerprint analyses, where newer biometric technologies provide an opportunity to automatically search the Dutch national police database that holds 2.4 million headshots of suspects. The reference to face comparison intends to separate Catch from other facial-recognition systems and emphasise that experts and not machines have the decision-making power.

> I really see it as a mere selection tool and not as a facial-recognition system, that is why we explicitly say face comparisons, as the decisions are made by trained experts. (lead biometrics specialist)

This distinction in which the role of technology is limited to making an initial selection, narrowing down the number of possible matches within a large database, while the decision-making power rests with the trained expert allows practitioners to centre the discussion on the processes and procedures that govern it use, rather than the technology itself.

In the context of forensic investigation, the practitioner argues the process is designed to ‘minimise the error rate and ensure that the introduction of bias is minimised as far as possible’. Here, bias is defined as an overarching concern – it’s about people and their cognitive processes, which requires the police to organise the processes in such a way that it minimises error rates. Drawing on their forensic experience with fingerprinting, the biometric centre has split the process into a number of clearly demarcated parts. It starts when the biometric unit receives an investigative photo. They first check the quality of the image and the investigative question: is it a 1:1 comparison for authentication purposes, to see whether this person is who they say they are? Or is it a 1:N comparison for identification purposes, matching an unknown person to a known identity? The letter N is used to describe the sample against which the unknown face is compared. Catch cross-references this image against the police database to produce a shortlist of possible candidates. The next step is for two experts, independent from each other, to systematically check the shortlist for differences and similarities and judge the extent to which the photos are a possible match. Then the

forensic department ‘adds those two conclusions together and report the most conservative one’, as ‘we just do not know yet how accurate it is, this is still very insecure’ (lead biometrics specialist). These reflections point to a hesitance within the police to rely on biometric recognition for forensic evidence analysis, as the technology is relatively new and still shrouded in unknown unknowns. Thus, in the Dutch forensic context, biometric recognition systems by themselves are not seen to produce conclusive evidence about a person but are considered a small part of the identification process. Its function is to automatically sieve through a database for a 1:1 or 1:N comparison and compile a shortlist of possible matches in relation to a crime that has occurred. The decision-making power is presumed to stay with trained experts.

Projecting into the future, the increased volume of data that is publicly available on social media data infrastructures are believed to change the nature of policing. It offers the potential for police to capture evidence on a specific incident that are documented by third parties.

> The images that are available on the internet – before the police even realise there are violent riots, so to speak, the internet is already littered with images of it. So I do see, I’m also setting up my organisation for that, that we’re going to get more and more of that kind of request and that we’re going to have to invest heavily in it. (lead biometrics specialist)

This changes the spatial-temporal nature of what constitutes a crime scene within the context of forensic investigation, transforming police evidence collection from a physical space where the police control what and how data is collected to a more diffuse notion of a crime scene, where evidence is also created, collected, and shared by others.

6.1.2 Digitale Perimeter – the Netherlands

In the lead-up to the European Championships in 2021, the Dutch national police, the city of Amsterdam, the Johan Cruijff Arena,[14] and the Netherlands organisation for applied scientific research joined forces to develop a ‘digital fence’ in the area around the football stadium, called the ‘Digitale Perimeter’ (Amsterdam, 2021). The football stadium is demarcated as a ‘living lab’, a name given to areas in Amsterdam where different stakeholders can test technologies in the wild. The ‘Digitale Perimeter’ is an umbrella term for a range of technologies tested in this area, but in

[14]The Johan Cruijff Arena is the main football stadium in Amsterdam and is located in the Bijlmer, a migrant neighbourhood, and is part of a larger shopping district.

this chapter, I will specifically focus on the privacy-by-design face comparison system that is being developed by the Dutch police.

Technically, the proof of concept for the privacy-by-design face comparison technology relates to developing a system that allows for edge processing, a decentralised data process that abstracts and calculates a 'faceprint' on the end device, the camera. Here, a faceprint refers to the statistical calculation of the distance between certain facial features, like eyes, nose, and mouth, to create a unique ID (Introna and Wood, 2004, p. 186). Calculating a faceprint on the edge and merely storing the statistical calculation and not the image is seen to limit possible privacy infringements on the 'innocent' masses. The chief innovation manager observes:

> All kinds of technologies that you can apply that in itself are intrusive to privacy, the puzzle is can you make applications of that that are less intrusive or that are more specific?

Privacy-by-design solutions in the context of policing are important: 'I have a lot of data that is none of my business and second, we as the state, but not just the state, are not very good at protecting this data'. He signals that there is an inherent conflict in the desire of the police to collect data, ensuring the privacy of the masses, and the organisational ability to keep this data safe. This experiment allows the police to collect data in such a way that what they are left with is a

> fairly abstract description of the face in the form of a number of ratio numbers [faceprint], and those ratio numbers can, of course, be linked to a ticket number. (chief innovation manager)

In theory, this proof of concept is designed to optimise for privacy; once a picture is taken, a faceprint is calculated and the image is discarded. The practitioner explains its intended use: in the event of a violent altercation or another disturbance, the face that is captured on the stadium's security cameras is used for a 1:N comparison. It compares ratio numbers to ratio numbers connected to a ticket number, which is connected to a name, a bank account, and other personal identifiers. This type of data processing should allow police to narrow down the search from all people attending a match or concert to a small group of visitors and points to two developments. The calculation of faceprints makes the football stadium a site of a future crime scene, where police are allowed to pre-emptively collect data about who enters into a demarcated space to facilitate future yet unknown investigations. It also moves the identification process beyond a 1:1 or 1:N comparison of an unknown face against a police database of known identity to one in which the

ratio numbers of a face become the entry points into third-party databases, in this case, financial data infrastructures.

6.1.3 Zaventum airport – Belgium

In the aftermath of the 2016 terror attacks, where two suicide bombs detonated at the Zaventum airport near Brussels, the Belgium federal police started a facial-recognition pilot at the airport, which ran a little over a year. Images collected from the surveillance cameras in the departure area were compared to a synthetic blacklist. The policy adviser to the general commissioner of the federal police[15] explained that the synthetic blacklist consisted of 'about twenty most-wanted criminals', supplemented with airport staff who participated in the pilot. The aim was to see whether the system would filter out people on 'the watch list from the sea of passengers'. This notion of cross-referencing a synthetic blacklist to individuals moving in a demarcated space indicates that the face is used for a 1:N comparison, with the objective of identifying a known suspect in a mass of people in real time. The system was not running or monitored continuously but rather sporadically:

> not operational full-time. Concretely, there was a PC inside the airport police office, and every now and then, someone would effectively sit behind it to see if any hits came in. (policy adviser)

When it was turned on, the police experienced high error rates. 'I've seen examples of a young blonde lady was matched with a passenger, a man in his sixties with glasses and a moustache' (policy adviser). These experiences were not what ended the pilot, however the trial was halted by order of the Belgium Supervisory Body for Police Information Management, which launched an investigation after they learned of the pilot through a press statement released by the police. The member-adviser of this body explained that the issue was not about the processing of biometric data in itself, which is legally allowed in specific situations, but there are restrictions on constructing technical databases.

> One of those are the technical databanks.[16] What differentiates it from others is that it involves the automated collection of data, where basically in the collection there is no human intervention, which is only allowed for ANPR data from ANPR cameras.

[15]The commissioner general heads the Belgium federal police force. This office is responsible for the management, strategy, and policy of the federal police and the coordination with the local units. Their office is supported by four policy advisers.

[16]Police databases are referred to as databanks in Belgium.

Again, a distinction is made on the loci of decision-making power and which entity is delegated the final authority to act, not just at the endpoint, deciding whether there is a match between a face and an identity, but also in the information architecture, deciding how data ends up in a police database.

These experiences have not diminished the Belgium federal police's interest in facial recognition. In response to the halting of the pilot, the general commissioner's office erected a biometric working group that is assigned the task to identify the biometric recognition policing needs and the legislative obstacles that are currently preventing its use. Thus, the aim of this working group is to propose legislative changes that will allow for future uses, as the policy adviser who leads this working group explained: 'not technically but just policy-wise, what can we do now to start with facial recognition in the future?' (policy adviser) Looking to the future, there is a clear police desire to identify certain individuals, 'terrorists', in real time in public spaces and govern how they move about domestically.

> The desire is there, in the first place for terrorism – this is the crime type that we had in mind, so identifying persons in a crowd to prevent them from committing violent offences […] or other events in light of critical places where we want to control who accesses it. We now do this physically but we believe there are technical tools that can assist us in this. (policy adviser)

It is thus when this policy adviser talks about the possibilities of facial recognition in relation to terrorism that its affordances shift away from the current ability of technology, automating face searchers in databases, to future possibilities, the role technology can presumably play in pre-emptive identification and controlling those people who are classified as potentially dangerous or as showing suspicious behaviour in specific locations in real time.

6.1.4 Preparing for biometric recognition – Denmark

The Danish federal police are exploring the possibilities of facial recognition for policing purposes; as such, I will not present a specific tool here but their process of preparing for the introduction of biometric recognition technologies. The chief privacy officer of the federal police start by explaining its current use cases. In the context of airport security, a traveller scans their passport upon entry or exit, a 1:1 comparison for the purpose of authentication, and in the context of child sexual exploitation, there is an experiment to try and make seized materials searchable. He quickly adds that the experiment is not deployed on a real case.

> Now we are not formally allowed to explore, as we actually planned that we would, a project that explores the use cases and the risks of using facial recognition across different areas. Because it is simply deemed too much of a hot potato at the moment. (chief privacy officer)

Plans to explore different facial recognition uses are contingent on the political climate towards the use of digital biometrics by police. He continues to explain that 'it will definitely come up again in the fall', demonstrating that there is reason to believe that political controversy related to facial recognition will only temporarily restrain its use. Looking ahead, the chief privacy officer argues that, to navigate the political climate, address policing needs, and find the best way to utilise this technology, they will most likely select one contained and non-controversial use case, where the police will be 'conservative in what we would propose'. This entails choosing a technology that is 'relatively easy and inexpensive to deploy'.

The Danish process is characterised by finding a balance between the desire of the police to use these technologies and the due diligence in weighing the risk and actual possibilities of biometric recognition by including a variety of perspectives, which should allow them to move beyond the sales pitches.

> A full 360 review, so we want to look at the risk, look at the technology, talk to external experts, academia, explore myths and assumptions. Sometimes vendors of products in this area might be making assertions of how useful this is, [and] we want to make sure it actually is. (chief privacy officer)

This should inform them where there is an 'actual business case where this makes sense to police' (chief privacy officer). Throughout the interview, he emphasised that the police is a professional organisation that carefully weighs different interests, has proper due diligence processes, and is aware that sales pitches of computational vendors might overpromise the accuracy and usability of the technology. Even if these technologies lived up to the vendors' promises, they might still not be sufficient for an organisational policing context, which has to balance the actual benefits of technology for policing and the resources required to run a these system within the organisation.

6.1.5 Speaker identification Integrated Project – International

Voice identification is a more nascent and unknown field of biometric recognition technologies. Here, I will explore SIIP, a research project supported by the European Union's funding programme

FP7 and awarded to a consortium of a military-industrial company, police forces, and technology and university stakeholders.[17] The participating policing organisations are Interpol, the Bundeskriminalamt, the London Metropolitan Police, the Ministero della Difesa (Italy), and the Ministério da Justiça (Portugal). The project aimed to build the first international and interoperable database of voice biometrics to support investigations in the context of transnational threats, in other words, terrorism and organised crime, and builds on the assumption that one of the most prominent contemporary obstacles in the fight against organised crime and terrorism is the fact that suspects can use multiple and arbitrary identities (European Commission, 2018). As a practitioner outlines, 'law enforcement agencies at times are clueless when there are no fingerprints involved' (programme manager), and explains:

> At the time, we started to see more appearances of groups like ISIS publishing videos on YouTube, where they were recording executions of prisoners, and in some of them, their face was covered. So there was a need to work on the basis of the voice only. (data protection officer)

This observation suggests that the fact that police are at times confronted with high-profile violent incidents, where voice is the only investigative lead, demands a response from them.

Technically, SIIP offers police a number of new features. It has developed the ability to match unknown voice samples of a suspect to a corpus of data existing in police databases or collected from data infrastructures such as social media platforms or phone infrastructures. As a practitioner observed, now 'crime scene means any place or any investigation where police can record a person' (programme manager), pointing to the expanding spatio-temporal nature of evidence collection. What is new about SIIP is that it allows the police to construct a biometric profile through a process of attributing demographic characteristics and personal traits to a voice sample, what is also referred to as 'soft biometrics' (Abdelwhab and Viriri, 2018; Dantcheva et al., 2015; Kak, 2020).

> The unique thing about the project is that we were using seven different engines to try and find a criminal. Seven engines mean keyword spotting, age identification, language identification, gender identification. (programme manager)

[17]The main stakeholders are Verint Systems Ltd, an Israeli military and security company, which is coordinator of the project; Nuance Communication (formerly known as Loquenco); SAIL LABS; Idiap, which provided the technical capacity; and Interpol, the Bundeskriminalamt, the London Metropolitan Police, the Ministero della Difesa, and the Ministério da Justiça were the participating law enforcement agencies.

The project as such moves from comparison (1:N), are these faces the same, to attributing specific meaning to a voice, such as age, gender, and accent. The affordance of a soft biometric profile lies in its perceived ability to function as a preselection mechanism, narrowing down the total volume of data available on social media to those voices who share a soft biometric profile with the perpetrator (M:N selection). The letter M describes the data available on social media from which an N sample (voice reference database) is selected.

The research part of SIIP has since been completed, the technology is integrated into Interpol's biometric databases (Kofman, 2018), and the consortium started working on a research project called Roxanne. Roxanne integrates different technologies such as voice recognition and language and video technologies for the purpose of social network analysis. 'This project is about identifying organised crime. SIIP was about finding one unknown speaker, and Roxanne is about identifying this whole network of criminals' (programme manager). Here, he draws our attention to an N:N comparison, creating relationships between audio and video samples to gain an understanding of who is associating with who.

I will conclude my description of biometric recognition uses by summarising my findings. Looking across and between these five biometric recognition technologies allows me to explain the current state of play and the possible futures of data-driven policing. Each of these projects is in different development stages, Catch has become embedded within the forensic department, and the other projects are either still in development, in a state of reconfiguration, or have materialised as a database and a next-generation research project. Still, all of them speak to the belief that abstracting bodily features, such as face and voice, will allow the police to engage in a range of activities:

- 1:1 comparison for authentication purposes, to test whether this person is who they say they are;
- 1:N comparison for identification purposes, to match an unknown person to known identities;
- M:N preselection mechanism, to narrow down the total volume of data available on social media to a voice reference database of voices who share biometric characteristics with a biometric profile;
- N:N comparison, to create relationships between samples to gain an understanding of who is associating with who.

Embedded within these activities is the promise that biometric recognition will allow the police to address certain policing needs, such as automating existing tasks and being able to access and use the volume of data that is now available in a wide range of data infrastructures, such as police databases, social media platforms, phone infrastructures, and CCTV cameras. The digital abstraction from the face and voice becomes a unique identifier that can tie together information about a person across time and space. As such, it comes with the promise that this function will allow the police to expand the nature of evidence collection from analysing confiscated material objects from a specific location after a crime has occurred to the collection of material and digital artefacts from both the crime scene and beyond. When practitioners look towards future biometric recognition, they express an interest in identifying and monitoring individuals across time and space for the purpose of investigation and control. The constantly evolving and expanding practice of biometric recognition demonstrates how it is important to explore how the logic of this function is shaping how police come to understand and act on crime. Therefore, in the next section, I will explore the different meanings inscribed into biometric recognition technologies.

6.2 Recognition: An external organisational optimisation logic

The scholarly debates discussed in chapter 2 situate biometric recognition as a central feature of contemporary governance that allows corporate and state actors to more reliably tie a single stable identity to a person (Leese, 2020). The concerns are that these technologies perform less well on certain demographics, pose challenges to our fundamental rights, and that deploying them in real time creates a presumption to intervene (Buolamwini and Gebru, 2018; Fussey and Murray, 2019; Kak, 2020; Kind, 2019). These critiques foreground important issues, primarily related to facial recognition, through engaging with the technology as an isolated artefact, a practice that needs to be regulated and a function that solicits a certain response from front-line officers. However, as I will come to argue in this section, they do not fully capture the meanings practitioners ascribe to biometric recognition as a technology of governance. In this next section, I will engage with the different notions of recognition that emerge from my empirical research. Here, recognition is a construct for automating identification, attributing specific meaning to an identity, justifying the expanding nature and scope of the police, and affirming police as a competent authority. Although I engage with these ideas as isolated constructs to highlight their specific affordances, they should not be seen in isolation from each other but rather as a multitude of meanings that are inscribed alongside and in relation to each other. I will conclude this section by arguing that these different notions of recognition point to an underlying external organisational logic, in which biometric

recognition is understood as a construct that allows the police to negotiate their position in society in relation to the state and the public at large.

6.2.1 Automating identification

Almost all practitioners argue that biometric recognition will support police in automating existing processes and position it is as a natural progression in the identification processes in the context of forensics and investigation. Take, for example, the data protection officer's response to why police were interested in voice as a technology for recognition:

> The organisation was already working on biometrics, especially facial recognition at the time, so it was also very interesting to develop further the research and the work on voice recognition.

For the police partners in the SIIP project, voice identification feels like a natural step from the digital biometric technologies that are already in use. Another argument that is often foregrounded is that identification is the heart of policing, and biometric recognition technologies can automate and speed up existing labour-intensive practices, which opens up police investigation resources.

> We saw it as a chance. Almost every investigation starts of course with the identity question: who is this, who did it, who committed this crime, who is the victim, who left this trail of evidence. We said, we have more like 2.4 million photos of suspects, and face comparison is nothing more than a big search engine that allows us to efficiently search through these images. (lead biometrics specialist)

The practitioner continues to explain how, prior to the deployment of Catch, police would manually search through the database, ask their colleagues, and at times even post an image on national television to identify an unknown suspect. Thus, automation is attributed to the possibility of more instant and demarcated processing of large volumes of data. This notion of automation was a recurring theme in my interviews.

> If you could utilise an algorithm to look through that in an efficient way rather than having a police investigator or many police investigators spend hours and hours looking at tapes then that would be a real benefit. (chief privacy officer)

There are two affordances attributed to the automation of identification processes: increased speed and accuracy. To gain a sense of the velocity of data analysis, I will use a comment by an expert at the Interpol conference on SIIP, who explains that ‘a machine does about twenty thousand comparisons in around five seconds. This can never be reached nor validated by experts’ (Interpol, 2018). Here, the affordance of speed is understood as a binary between machine and man. Similarly, accuracy is also presented through the binary; the Danish chief privacy officer explains that watching camera images is mind-numbing and lengthy work that requires high levels of concentration and regular breaks, and even then it’s not foolproof; in contrast, biometric recognition tools don’t get tired. Automating identification thus speaks to the managerial logic of increased efficiency and effectiveness, as operationally biometric recognition technologies are believed to enable police to better process large volumes of data.

6.2.2 Attributing meaning to an identity

The construct of recognition also enables the attribution of meaning to an individual. Newer biometric recognition technologies are presented as a way to identify a suspect when there is little other evidence. Attributing meaning, such as age, gender, and accent to create a soft biometric profile allows for a new kind of recognition. This is best explained in the context of SIIP, as it offers a next-generation biometric recognition system.

> So in open-source intelligence, you can use all these machines to, for example, narrow down space where you are looking for a speaker. Say you are looking for a speaker that speaks Arabic, is male, and is an adult, but he has an accent […] like Saudi Arabia, you could narrow down the search space into these features and hopefully, you can narrow it down and maybe the chance is better to really find this fish in the sea. (Interpol, 2018)

A close reading of the project documentation further reveals that this soft biometric profile makes voice a search criterion for social media, a filter to select only those voices from YouTube that share voice features with an unknown perpetrator in the context of terrorists (European Commission, 2018). Thus, the soft biometric profile can ‘actively generate recognizability’ (Amoore, 2011, p. 69) by inscribing meaning to the voices of those individuals and communities who share demographic characteristics with an unknown perpetrator of a violent criminal offence. The system is actively trained to identify specific accents, languages, and genders, which makes certain individuals and communities more visible to the police.

A search criteria meaning is also attributed to voice to distinguish between known and unknown voices. In the context of organised crime, the affordances of targeted data collection on the basis of a biometric profile should enable the police to exclude the voices of innocent family members, whose voices are included in audio samples from phone taps, from further data processing. As a practitioner outlines:

> A platform where we were combining all of these elements and to try and find out the identity of the criminals, so there are chances of finding a suspect and to identify the innocent as well. (programme manager)

Thus, it is believed that soft biometric profiles allow police to account for some of the externalities of surveillance. Filtering out the voices of innocent family members should prevent their voices from being stored and processed in relation to a criminal investigation. This begs the question of who decides what meaning is attributed to a voice and how the terms ‘terrorist’ and ‘innocent’ are defined. The technical process is the same: soft biometric features are extracted and computed to create a profile. The difference is that the voices of innocent people who share biometric characteristics with a terror suspect are not labelled as ‘innocent’ but rather as a ‘sample’ to narrow down the search field. Here, it is the police, not the technology, who determine how someone is recognised.

Another way in which the attribution of meaning materialises in relation to biometric recognition was foregrounded in my interviews in relation to Catch.

> Then you have an image of that person, who has first swiped the pin code from an elderly person, then rolls their card and is standing at the ATM. (lead biometrics specialist)

Again, there is a situation where the only evidence on an unknown perpetrator is the surveillance footage taken from an ATM. The criminal act that was given serves as an example to justify the policing need for Cash, and similarly to the justification of SIIP, it illustrates that meaning is not only attributed to biometric features but also to a crime category. Here, the image of an ‘immoral’ perpetrator that preys on vulnerable senior citizens, by shouldering them at the ATM to uncover their pin code, rob them of their bank card, and withdraw their money, is a justification for why the police need biometric recognition technologies. The meaning that is attributed to this crime category plays into people’s emotions and justifies its use.

6.2.3 Justifying the expanding nature and scope of police

So far, I have discussed how practitioners inscribe the affordance of automation and profiling on the construct of biometric recognition. In this section, I will build on the findings at the start of this chapter, where specific projects were seen to justify the desire to expand the scope of evidence collection and its use offer biometric futures in which technology can aid the police to monitor and control how certain people move across a specific geographical area. Here, I will explore how these self-articulated policing needs justify the use of biometric recognition tools and offer a political rationale for why and how police need to expand their nature and scope.

A practitioner observed how off-the-shelf tools used in other policing contexts pique the interest of police in the office.

> If you look at Clearview[18] and the discussion behind it, you can say what you want, but you could solve issues with it, of course, it is ethically totally wrong and privacy-wise even more so, but the need is there, of course. The need is enormous. (lead biometrics specialist)

The desire to work with facial-recognition tools is conflated with the actual need for its use. It has to be noted that the forensic department does not use Clearview; rather, the lead biometrics specialist used this example to demonstrate how the existence of technologies creates a demand for their use. What is more common is the observation that police are confronted with crimes where the only investigation lead is an image of a face or a sample of a voice of an unknown perpetrator.

> In some crime, you only have voice as evidence to basically identify the unknown criminal. I think if we can see the example of the Charlie Hebdo [attack], where voice was the only evidence for law enforcement agencies and with some speaker identification technologies, they found a guy who was a singer or rapper on YouTube. (programme manager)

A very high-profile violent offence is described to demonstrate a clear investigation need, a situation that demanded action from the police and justified the use of voice recognition and data collection from social media platforms.

[18]Clearview is an American facial recognition company that has sparked global controversy on indexing billions of faces from photos scrapped of the internet and overstating the effectiveness of its product (Statt, 2020).

Times of crisis, specifically high-profile violent crimes that instil a collective fear in the public such as terrorist attacks, legitimise the need for digital biometrics and accelerate its deployment, as police feel a sense of urgency to respond or show the world they are responding. The lead biometrics specialist of the Dutch forensic department explains that, during the pilot phase of Catch, police intelligence suggested that one of the terror suspects of the 2016 Belgium attacks was hiding in the asylum circuit in the Netherlands.

> At the time, police management decided that, in light of national security and the terrorist threat level, we needed to implement a facial-recognition system within the national police force. This was a key trigger, an accelerator; eventually, we would have gotten it, but then, like in any big company, it would have ended up on the IT agenda and I dare say that we would not have had it by now. Now, I literally got a letter from the deputy chief of police telling me to purchase and deploy the system. (lead biometrics specialist)

The practice of data-driven policing is thus directly shaped by crisis – here, the desire of police leadership to act in times of crisis in combination with the perceived affordances of technology, in other words, to improve identification of terror suspects within the asylum circuit, created an environment in which the deployment of technology is accelerated. The details of how the deputy chief of police imagined Catch would identify terrors suspect is unclear, but it points to the notion that times of crisis justify and accelerate the police's interest in biometric recognition technologies.

I'll end this section of recognition as a construct to justify the expanding nature and scope of police with the following quote:

> If we are trusted to use it in the fight against terrorism, and we then start scaling it up in the shopping streets, then we are damaging that trust on our side as well. Trust is based on clear agreements and transparency. (policy adviser)

This reflection is meant to offer reassurance that the police is aware of the possibility of scope creep when technologies become embedded within the organisation, but rather offers insight into the crime categories on which police believe more intrusive biometric recognition systems can be deployed without damaging trust and confidence in the police.

6.2.4 Affirming police as a competent authority

Times of crisis also point to the final meaning practitioners inscribed onto recognition, that its use reflects on the professionalism and competency of the police as a legitimate authority. The 2016 Brussels terror attacks prompted the Belgium federal police to pilot facial recognition on the airport of Zaventum. During the pilot phase, the federal police released a press statement:

> To say, look, the police are really taking additional efforts to prevent something like this from happening again. (policy adviser)

In this sense, the police have a desire to communicate about their uptake of biometric recognition technologies to show the public they take action after a high-profile crisis. Here, the mere act of experimenting with facial recognition, not its actual usefulness in identifying terror suspects, is linked to increasing positive public sentiment towards the police. This speaks to the concept of surveillance theatre (Van Brakel, 2021b), which finds its origin in the idea of security theatre, where visible security measures aim to provide a sense of security, but in reality, provide little or no actual security (Johnston and Warner, 2010; Schneier, 2003). Still, when we situate this observation beyond the initial response, which I will come to argue in this final section on meaning inscribed to recognition, the (visible) turn to biometric recognition technologies is primarily aimed to reinforce the notion that police are a legitimate and competent authority in European societies.

Several practitioners indicated that they feel it is part of the police mandate to develop and deploy biometric recognition technologies.

> So what I find problematic is that we act as if the testing and development of this kind of technology would be better left to other entities than the government. What I find strange is that the whole debate is about the fact that the police are looking at it. I would find it very weird if the police were not looking at it. (programme manager)

Here, the programme manager ties the testing of biometric recognition technologies to the mandate society gave to the government, and by extension, to the police. The feeling that police is the legitimate authority to look at and test this function is reinforced by the external pressures experienced by practitioners. The innovation manager explains:

> I think the FR technology is something that exists; it is coming at the police and we can't avoid it. This is one of those technologies where, in certain situations, if I could prevent a serious crime, you will be blamed for not applying it. (innovation manager)

Practitioners feel a certain inevitability of its use, that because the technology exists, the police are expected to use it.

When looking across and between the different policing contexts, it becomes clear that, when it comes to biometric recognition technologies, the police have a complicated relationship with the public. Police feel that parts of society will judge them for not using all possible tools available to them to prevent, manage, and solve crimes. At the same time, they feel they are being judged for its use, as the turn to biometric recognition has been subjected to increased levels of public scrutiny. One practitioner said that it is inevitable for someone to say:

> How dumb is it that I can do these things on my phone but that the police cannot do it. Why are you, the police and government at large, such a bunch of idiots? (programme manager)

> I do understand the difficulty; society is vocal about the police violation of our privacy, but on the other hand, society is also asking for criminals to be caught and spotted and for us to prevent an unacceptable attack on personal integrity or a threat to a large mass of people. (policy adviser)

The dynamic in which different publics have different expectations of the police use of biometric recognition is creating friction makes some police officers less inclined to discuss the topic outside of the organisation and others more apprehensive to use it, as in the case of Denmark. Police practitioners in the UK, a jurisdiction where remote facial recognition has led to much public controversy, declined requests for interviews or merely wanted to discuss the matters off the record. European police practitioners in less politicised environments mentioned that

> there is a lot of stuff happening internationally with facial recognition right now, [and] there is a lot of talk about artificial intelligence internationally, so it [the organisational development] is a bit in limbo. (chief privacy officer)

> Of course, it impacts us when a foreign police department arrests someone on the basis of an unverified score from a system. (lead biometrics specialist)

The use of biometric recognition has become politicised, where its use and non-use is scrutinised, and public sentiment, in part, is influenced by events that happen in other jurisdictions.

Practitioners feel that the debates surrounding biometric recognition systems are skewed on a number of accounts. First, in my off-the-record interview, a high-level police practitioner argue that there is an imbalance in the public debate between the interest of those who are confronted with the negative impacts of a crime and those who might be wrongly accused of committing a crime, arguing that scholarly and public debates are currently privileging who these technologies might disadvantage over those it could benefit. On top of this, there is a feeling that the debate is misinformed.

> The whole racial bias story is based on wrong assumptions and misinterpretation of studies. [On] vendors that, let's just say, don't deliver real production systems. (lead biometrics specialist)

Here, the lead biometrics specialist refers to his observation on how critiques are conflating systems that are not suitable for policing operations with systems that are actually used. He continues to explain that, while racial bias in facial recognition was not a point of attention at the time, Idema[19] was selected as supplier of Catch, and it scored as very reliable in the recent National Institute of Standards and Technology (NIST) test. 'Fortunately, we see that, at our supplier, it is not so bad. In fact, it is almost undetectable' (lead biometrics specialist). Practitioners refer to NIST as an authoritative voice that 'has conducted tests to quantify demographic differences for nearly 200 face recognition algorithms from nearly 100 developers' (NIST, 2020). The lead biometrics practitioner continues to argue that his unit works with those technological suppliers that are in the 'champions league' of forensic science and have extensive experience with the organisational context of the police, in other words, 'setting up a production system with such a database, with poor-quality images that have not been checked' (lead biometrics specialist). In this reflection, he situates the public controversy on biometric recognition technologies in relation to the levels of trust and confidence in the police as a professional, competent, and reliable organisation. Where contrary, to public opinion, the police know how to select good technological suppliers and recognise the weaknesses within their own data infrastructure.

[19]Idema is the computational vendor of Catch.

The connection between a certain action and the trust and confidence in the police as an authoritative organisation is made even more explicit in the reflection of the programme manager of the Dutch police:

> Police is the enforcement association of citizens of the Netherlands. And then you do that in a certain way, you may also trust that it is done professionally. (programme manager)

The programme manager directly relates the justification of biometric policing practices to the police mandate and claim to power, painting a picture of the police as a professional organisation that has been entrusted to perform a certain task on behalf of the people and, as such, should be trusted to do this to their best ability possible. This picture of the police as the association of all citizens that is a professional organisation is what Mulcahy (2013) refers to as the legitimacy process through representation, in which police communicate a particular image of policing to normalise specific actions.

I will conclude this section on the meanings that are inscribed onto the construct of recognition with the reflection that studying data as practice allows me to situate the emergence of biometric recognition technologies within the operations of policing. My empirical findings provide evidence for four notions of recognition: a construct to automate identification, attribute meaning to an identity, justify the expanding nature and scope of the police, and affirm police as a competent authority. Contrary to my case study on data-driven risk scoring, practitioners who work on biometric recognition are aware of the external dynamics that surround this function. Therefore, in this concluding section, I argue that the advent of facial recognition and voice identification requires a response from the police, as practitioners directly relate its use and non-use to the public's perception of their professionalism and competency as a legitimate authority. This productive nature of biometric recognition is a dynamic that I will refer to as an *external organisational logic*.

6.3 Recognition as a driver of police practices

This chapter has thus far foregrounded that the turn to biometric recognition systems is seen as a natural progression from existing practices and practitioners feel that the mere existence of these technologies places normative expectations on the police, which make its use seemingly inevitable. As such, there are multiple meanings ascribed to the use of biometric recognition, which creates a

dynamic of an external organisational optimisation logic. These insights on to what end the police turn to biometric recognition contribute to my research questions, as they sheds light on the actual nature of data-driven policing and how it has defined how police come to understand crime and police power. In this next section, I will show how practitioners are pre-empting internal and external conflict surrounding the turn to facial recognition and voice identification by erecting a number of practices – that of pilots, transparency, and safeguarding.

6.3.1 The practice of pilots, living labs, and experiments

A common theme across the five use cases presented in the first section of this chapter is the language around pilots, living labs, and experiments. Catch and the trial in Zaventum airport started as pilots, the privacy-by-design implementation near the Amsterdam football stadium is part of something that is called a 'living lab', and SIIP is referenced to as a research project. These are three names for the same practice, that of testing technology within a demarcated spatio-temporal environment, at times 'in the wild'. In this section, I will explore the practice of pilots, their use, purpose, and discourse, to conclude that pilots are both a practice in themselves and a means to establish new practices.

The evolution of Catch shows a clear progression from a research phase, in which the police look at what these new technologies potentially can and can't offer, to a pilot phase, where an idea is tested in practice, creating a proof of concept, to the implementation phase, in which its use becomes embedded within the organisation. The lead biometrics specialist explained that they had been following the developments in facial recognition technology for a while and that 'at the end of 2016, after doing a pilot and writing a business case, we put that into operation within the police'. Here, pilots are seen as a crucial step in the introduction of technology in the police, as it allows them to move from the theoretical to the practical and gain organisational buy-in for its use.

> We can explain what we want to do, why, and under which conditions, after which we will lose their interest in ten minutes, but if I make it and people can touch it, they get a feel for it. (chief innovation manager)

Pilots are an internal vehicle to navigate the limited interest of senior management in technology and to promote the use of digital biometric within the police. External opportunities allow practitioners to create momentum behind a specific technology. In the context of the Digitale Perimeter, the pilot environment created by the city of Amsterdam offered the chief innovation manager the opportunity to develop a privacy-by-design proof of concept.

> Off the bat, I will agree that facial recognition in relation to football matches is not proportionate, but for me, it's all about the experiment. We had to come up with a use case for the field labs, and I believe this to be a useful application that we can try out in this context and then apply somewhere else. (chief innovation manager)

The 'living lab' allowed the police to develop and test a biometric recognition system that is less about ensuring 'increased security' for the European Championships and more about an individual seizing the opportunity created by actors external to the police. The living lab emerged from the municipality's desire to test technology and created conditions in which police could retrofit specific ideas to the external opportunities presented to them. 'Having to come up with a use case' implies that the decision to test technologies in a demarcated space was made long before it was decided which technology would be deployed for which problem.

The Belgium police working group, established in the aftermath of halting the deployment of facial recognition in Zaventum airport, situate pilots as a mechanism to navigate the regulatory requirements imposed on the police. The police adviser observed:

> My concern all along has been that of the chicken-and-egg problem, do we need to have the legal framework before we can test it, only to find that perhaps that we don't support its use or that it doesn't meet our needs. Yes, then we have spent a lot of time creating a legal framework that serves no purpose. So I would rather turn it around. I would say give us a framework that allows the testing of it, under certain conditions, within a certain time frame, and then together weigh its appropriateness and whether or not legal adjustments are necessary. (policy adviser)

She draws our attention to the notion of regulated pilots that would allow police to overcome the current tensions surrounding the deployment of digital biometrics, the resources needed to create and comply with legal requirements, and the uncertain benefits for police. This view is contested. The oversight body that halted the Zaventum facial-recognition trial explained that there were a wide range of issues with this project, from proportionality and information security to accuracy issues. While the pilot was halted on the inadequate compliance with legal grounds that allow for the use of an automated database, there was more at stake:

> I would summarise it as amateurism. It was started in 2017 with a system that was absolutely not to the point, so there were all kinds of problems with it, but actually we

> stopped it for that reason as not to enter into further discussions. (Belgium data protection actor)

Here, two tensions emerge surrounding the practice of piloting. First, after a moment of crisis, the federal police choose to pilot a tool that was, according to an external review, not fit for a policing purpose. Second, when a technological system is halted, the oversight authority does not necessarily foreground all the issues involved, merely the one that is 'the easiest' to substantiate their argument and actions, in this case (temporarily) halt the practice. Primarily scrutinising data-driven policing on low-hanging fruit, a detail of the pilot, without engaging with more challenging or contested issues, the faulty implementation of the pilot itself, runs the risk of misinforming how the police can proceed with biometric recognition in the future.

The final observation I want to foreground is how practitioners are actively framing technology projects as pilots to make them less controversial. In the context of SIIP, practitioners stressed the research aspect:

> Just to let you know that these projects were research-based projects, so there was no such intention to run it through real data such as crime scene data. (programme manager)

> This was a research project, so at some point, there was a lot of work on what we can call research data, right. So that was the line we were using in the project. (data protection officer)

As such, labelling the development of technology as research is a conscious act that allows police to isolate biometric recognition systems from real-world implications and implementations. The frame of research is just that; since its inception, SIIP has become embedded within policing as Interpol's third biometric database (Kofman, 2018) and has prompted additional technological developments.

What we can learn from these reflections is that the discourse around pilots is deliberate and primarily seen as a vehicle to support the development and testing of technologies. Its practice creates demarcated space in which the police believe they can experiment and can move from the theoretical to the practical; it becomes a vehicle for new technology practice to emerge, as it allows the senior management to 'touch and feel' data in the hope to pique their interest; it offers a solution to navigate perceived regulatory constraints; and it is a frame to isolate the technology development process from real-world implications. As such, pilots can be seen as a practice in themselves, which

allows individuals in police to introduce something new, in this case, digital biometrics, but the process is also a means for new practices to emerge.

6.3.2 Transparency as a practice

Transparency as a safeguard emerged throughout my interviews as a tool to navigate external criticism. To pre-empt and curb public criticism, it is considered imperative to proactively explain the technology and its use, in other words, the choices and actions that emerge from deploying digital biometrics on the streets. Transparency, being open about how the technology is used, which data is collected, and how long it is stored for, is seen as a way to halt speculation that could decrease trust in the competency and authority of the police.

At the heart of the practice of transparency lies the belief that police are a legitimate power holder in society and resistance or public controversy is seen to stem from the lack of knowledge on what is happening and not the use of these technologies. The lead biometrics specialist reflects on the controversy relating to South Wales Police:

> They never explained properly what they were doing, and they accept that there are false positives, but that the only action they take is to check someone who they might otherwise have checked anyway. (lead biometrics specialist)

Similarly, another practitioner said:

> So I think above all we need to build trust in the way the technology works and the way it will be deployed, for which crime category, how we will act upon the results, and how we will deal with any mistakes. I think that's where the key lies to moving forwards in this story: working together on transparency. (policy adviser)

Most practitioners agreed that the use of biometric recognition speaks to a fear of constantly being monitored in public spaces. Therefore, to preserve public trust in the police, there is a need to explain what the police are doing, which comes with the challenges of explaining complicated technical processes in such a way the public understands it.

> How do you explain the use of facial recognition to the public, knowing there will be people who will not believe what I say because of my uniform? But the other problem is

> that these are conceptually difficult issues that you have to translate to something that the general public will understand. (chief innovation manager)

The effect of transparency is thus limited by the extent to which police can explain the use of biometric recognition and the audience that is listening, where some publics will always question the legitimacy of the police, and by extension, any of their practices.

The practice of transparency is a proposition made by the police that is responded to by different audiences and in some cases requires a response from them; as such, it is relational in nature. In this final section on transparency, I will show that this practice is stratified, as the responses of third parties require responses from different actors. In the Dutch context, the lead biometrics specialist reflects on the critiques on their use of facial comparison in the forensic context. When Catch was first introduced, there was resistance to its use; the local digital rights organisation made a public statement ‘we will make sure that it will be banned in the Netherlands’ (lead biometrics specialist), and parliamentarians asked critical questions about the implementation to the minister of safety and justice. This indicates that there are two distinct areas in which its use is challenged, in the public debate and within government structures. For the latter, he observed:

> Well, if you look at the minister’s answers to the parliament, the minister said, and I thought that was a big compliment, Catch is the only implementation where things are arranged in such a way that it provides sufficient safeguards. (lead biometrics specialist)

This reflection offers insight that the safeguard of transparency happens in different spaces: the police explaining to the public and the minister who is responsible for the police explaining it to parliament. ‘I do notice that the resistance to Catch has subsided a bit, and they are concentrating on other things’ (lead biometrics specialist). This suggests that practitioners understand public critique to be temporal in nature; it dies down when the technology becomes more normalised and other issues emerge that require public, civil society, and parliament attention.

6.3.3 The practice of safeguarding

So far, I have explored the practice of pilots and transparency that emerged from the use of biometric recognition systems. I will now turn to the practice of safeguarding, which I also discussed in my case study on data-driven risk scoring. In the context of biometric recognition, I will draw attention to the two types of safeguards that emerged: placing the loci of decision-making on the expert instead of the machine and building technology from a privacy-by-design perspective.

I will explore to what end each safeguard is it erected to conclude that these practices are aimed at mitigating a specific challenge the police face in their development and deployment of biometric recognition technologies.

The first practice was already alluded to in the first part of this chapter and relates to the loci of decision-making power, the trained expert or the machine. This dichotomy allows practitioners to situate biometric recognition systems as merely a selection tool that offers one or more choices for the expert to judge. To demonstrate this, I will reuse a quote from the start of my chapter:

> I really see it as a mere selection tool and not as a facial-recognition system, that is why we explicitly say face comparisons, as the decisions are made by trained experts. (lead biometrics specialist)

The practice relies on trained experts for the decision-making process related to the many unknowns that shroud this technology. In explaining why and how police turn to biometric recognition, its use is often positioned against the tried, tested, and extensively studied method of fingerprinting. Police hesitance to fully rely on biometric recognition stems from the idea that it has not yet stood the test of time.

> If you compare that to fingerprinting, we have been doing that for over 120 years; with faces we just don't know yet, that is still very uncertain. There is still much more scientific research that needs to be done on identical twins, if the face can be manipulated through plastic surgery, diseases, and other manifestations. (lead biometrics specialist)

Similarly, on voice identification, a practitioner observed that

> sometimes the normal voice is different than when you use the same voice at the crime scene, because it can give a different sound and different message. Or if you are waking up in the morning you have a different voice; it is a bit heavier than the normal voice. If the person is smoking or with the age you change your voice as well. (programme manager)

Here, the programme manager points to the fallibility of biometric recognition technologies, in which bodily features that can be extracted at a distance are not constant and can change over time

and can resemble others. Thus, the police feel they cannot blindly rely on the biometric output, which is where the trained expert come in.

> An aspect was not to give one result. When the machine provides you several results, we make sure that we provide at least three to five results that they can investigate. (programme manager)

So while there are still many unknowns surrounding biometric recognition, it can be used at the start of the production pipeline, creating a shortlist for experts. This safeguarding externalises the decision-making process, not on technology but on a trained expert.

The second safeguarding practice I will discuss is that of building technology from a privacy-by-design perspective. To understand how this practice emerged, I will first engage with what is driving the chief innovation manager. He situates his work within a vow every police officer makes 'to protect the citizens and their civil liberties, which include the right to privacy'. He continues to argue that

> it is always about security on the street, but we never question how unsafe is it? If you look at the list of the safest countries in the world, we [the Netherlands] are number five or six. So actually it is not that insecure. So the questions should be more about freedom than security. Of course, this is difficult to explain to someone who just experienced a crime – it is not like there are no issues – but we should be mindful of what we are giving up to make our society just a little safer. (chief innovation manager)

Dominant debates that instil fear in the public imagination will justify specific practices; in this case, the use of biometric identification technologies can control certain aspects of society, rather than having an honest discussion about balancing values of freedom, privacy, and security. This belief materialises in a privacy-by-design 'proof of concept' at the Amsterdam football stadium, which is argued to be less intrusive than other facial-recognition systems, as it calculates and stores faceprints on the edge. This is needed, as even with the best intention, it is important to acknowledge that the police is an operational organisation with the mandate to enforce the law and ensure safety and security on the streets; imminent threats or urgent cases create pressure on police officers to take action, which can't be constrained by procedural safeguards.

> I don't think it should depend on my good behaviour. Because in essence, I can't trust myself; I only have to be tempted once. Scope creep will occur the moment we know

> that we have a picture of some terrible terrorist – we just don't know which one. I can try and stop the use but that won't hold for very long. (chief innovation manager)

The existence of technological capabilities creates a sense of urgency in the development of privacy-by-design facial comparison. 'I think we need to make sure that we have an application that we believe in as soon as possible to prevent the people who want to make applications that are much more generic and blunt from taking the lead. So present an application of the technology that is less intrusive before the mainstream prevails' (chief innovation manager), pointing to a dynamic where digital biometrics will be introduced beyond forensic purposes regardless whether certain safeguards manifest or not.

To conclude this section, I will argue that the organisational optimisation logic creates the conditions for a number of practices to emerge – that of pilots, explanation, and safeguarding. All these practices relate to how the police can introduce this function with the least amount of friction possible. Pilots are a frame that allows police to carve out a space in which they can test biometric recognition technologies 'in the wild' and make abstract concepts tangible for senior management who are responsible for strategy and the allocation of budgets. The practice of transparency builds on the assumption that public and political controversy surrounding facial recognition relates to a lack of information and understanding rather than its actual use by police. Thus, if police can only explain their needs, the crime categories the technology would be applied on, and the processes and procedures that will govern its use, the deployment of biometric recognition will become less controversial. Responsible use is primarily defined as ensuring that the loci of decision-making lie with the trained expert instead of the machine and the use of a privacy-by-design perspective system that should prevent any possible misuse. Therefore, I conclude that, while biometric recognition technologies are often engaged with on a technical level or how it materialises on the street, these insights reveal that practitioners frame its use as inevitable, and it is to the police to find the best way to embed it within the organisation.

6.4 Conclusion

In this chapter, I explore the practice of biometric recognition to contribute to answering my research questions. What is the actual nature of data-driven policing? And what is the relationship between datafication and police power? The different cases foreground that police are actively engaging with newer biometric recognition systems to match an unknown suspect to a sample of known identities within the context of forensic science and investigation. When looking towards the

future, the perceived affordances of biometric recognition technologies create an idea that police will be able to use it for the monitoring and controlling of individuals across a territory. These perceived biometric futures expand the functionality from forensic to investigation and intelligence, in which more invasive forms of data collection, through social media platforms, CCTV cameras, and telecommunication infrastructures, and different types of analysis, from biometric comparison to soft biometric classification, are positioned as inevitable.

Looking across and between these specific implementations allows me to distil a number of insights. First, the turn to this function is driven by the belief that it will allow police to expand the nature and scope of evidence collection to include data created by others and data created by machines from a range of data infrastructures. Second, to what end police turn to biometric recognition is discussed in the meaning practitioners ascribe to its use, more specifically its affordance to automate identification processes, attribute meaning to an identity, justify the expanding nature and scope of the police, and affirm police as a competent authority. Particularly the latter two foreground how digital biometrics require police to actively engage with their external environment, both the public and the political landscape, to ensure that its use heightens rather than lessens the levels of trust and confidence in their professionalism and competency as a legitimate authority. I refer to this dynamic as an *external organisational logic*. The practices that emerge from this external optimisation process, that of pilots, explanation, and safeguarding, all relate to creating the conditions that allow biometric technologies to materialise. Practices are not directed at questioning the need or role of biometric recognition but at mitigating the negative externalities, challenging the professionalism and competency of police as a legitimate authority, that can emerge from its introduction. This indicates that the use and non-use of biometric recognition have broader implications for our understanding of police power.

7. Civil actors responses to data-driven policing

In the previous chapters, I have explored the use of data-driven technologies and I have shown that it is becoming a prominent feature of contemporary policing. These chapters offer insights into data as practice, in which I explain the origin, uses, and futures of data-driven risk scoring and biometric recognition to reveal to what end and on which grounds these tools become integrated within policing. While data-driven policing is often discussed as an attribute that is computed about an individual in relation to a crime category, what my research shows is that police believe that these functions will allow them to optimise for certain organisational needs and reaffirm their strength and professionalism in times of crisis. I name this dynamic the *organisational optimisation logic* of data-driven policing, where data impacts how needs, challenges, and tensions are defined from the perspective of the police officers and results in a number of new practices. However, as I discussed in chapter 2, social science debates on police power have theorised about its dialogic nature, where it is a proposition made by a power holder, which is responded to by different publics, which in some cases requires a response from the power holder. Looking at police practices will only partially answer my research questions on the nature of data-driven policing and the relationship between datafication and police power, and therefore requires an inquiry into its emergence as a site for struggle.

To explore data-driven policing as a site of struggle, I draw on Young's (2011) position that justice is political, 'rather than a top-down diagnosis of social life with a knowing initiator, a sense of justice in this context arises not from looking, but from listening' (Dencik et al., 2018). Therefore, to understand what is at stake when police power becomes embodied and enacted through data systems, this chapter draws on twenty interviews with civic actors to identify the injustices that emerge from the introduction of data-driven policing. Civic actors are defined as formal civil society organisations working on issues of digital rights, human rights, anti-discrimination, and racial justice, and individuals who work with and alongside racialised and impoverished communities who are subjected to police actions in Belgium, Brussels, the Netherlands, and the UK. This chapter is organised as follows. First, I outline the injustices raised by civic actors to conclude that there is not one single understanding of what is at stake when it comes to the introduction of data-driven policing. Rather, the range of civic responses to its emergence shows that there are three distinct entry points into the discussion on police use of data, what I refer to as the privacy, human rights, and race lenses. Second, I will bring together these, at times conflicting, lenses to highlight how data-driven policing speaks to a new politics of injustice in which its introduction is affecting whose voices and actions count in the discussion on police power. I

conclude that the introduction of data systems in policing creates competing priorities of what is needed to create a just society, which has broader implications for how we come to understand social justice concerns and police power.

7.1 The injustices of data-driven policing

This chapter will start by exploring the concerns raised by civic actors on the increased interest in and use of data-driven policing technologies in Belgium, the Netherlands, and the UK. The scholarly debates discussed in chapter 2 foregrounded that the introduction of data-driven policing functions runs the risk of perpetuating and intensifying the over-policing of communities of colour, creating broader privacy infringements, and overestimating the credibility of the system and thus acting more quickly upon its outcome. While civic actors raised similar and other concerns, their viewpoints offered a more nuanced understanding of what it means to introduce data systems within the bureaucracy of the European welfare state. I will structure these findings along five injustice claims: discrimination, the criminal justice trap, data protection, governance, and access to justice. Although I present them as distinct concerns to reflect how civic actors perceive what is at stake, they at times operate alongside and in connection to each other. It is not so much the specificities of the data-driven policing functions but the positionality, experiences, and belief systems of civic actors that affect what is considered the root cause of injustices – technology as a form of algorithmic governance, the state as the primary granter of rights, and the police as an oppressor of racialised and impoverished communities. These observations allow me to conclude that, while all civic actors position the introduction of data-driven policing in relation to police power, there is not one single understanding of what is at stake.

7.1.1 Discriminatory effect

The one common concern that materialises throughout all my interviews is that of discrimination, yet my findings show that civic actors ascribe multiple meanings to this injustice claim. Those civic actors who articulate their concerns in relation to a specific function or the affordances of data argue that these have a discriminatory effect and, as such, are unlawful, while civic actors who centre race in their analysis argue that discrimination is structural in nature and data-driven policing is just another tool in the historic and ongoing racialisation of crime. I will start by engaging with the observation that emerged from those civic actors who challenge discrimination within data-driven policing functions.

> The project is targeting mobile bandits who are defined by the police as pickpockets and thieves from Eastern Europe. So there is clear discrimination or distinction based on nationality that becomes a proxy for ethnicity. (human rights advocate)

She refers to the Dutch predictive policing pilot 'Sensing' (Amnesty International, 2020), which aims to identify suspicious cars within the ANPR systems and shortlist them for an identity check. The police automatically search for a predefined risk profile, attributed to an object, in this case a car, in a data infrastructure to control which individuals can freely enter a demarcated area. The discriminatory element is that this project optimises for the identification of a predefined racialised notion of a perpetrator, Eastern European and more specifically Roma, Sinti, and Travellers. A Dutch parliamentarian reflects on this relationship between pre-emption and discrimination:

> You combine data in such a way that certain groups become extra suspicious or receive additional attention, and actually it's a bit like stop and search. You start looking for other elements than for an actual identity of a person and end up including someone on the basis of his group membership. That is the actual definition of discrimination. (parliamentarian)

His observation reflects a well-known scholarly critique, discussed in chapter 2, that connects discrimination to the logic of prediction and pre-emption, in which reliance on police data to inscribe meaning onto individuals runs the risk of singling out a person on the basis of their group membership, in this case nationality and ethnicity, which further directs police attention to those communities that are already over-policed.

Social and racial justice actors offer a more complex and nuanced approach to the injustice of discrimination. From their perspective, racialised notions of crime are historically and socially determined and create visible and invisible, and immediate and systemic harms.

> Policing was designed to protect the upper classes from the masses, to protect their property, and fundamentally, it is still about policing poor people. That doesn't change, but what has changed is that its feels like there is an acknowledgement for the need for more inclusive, fair, and effective policing, but actually, the use of technology is just another tool in police's armoury and will continue the same patterns of entrenched discrimination that we've seen. (racial justice advocate)

Similarly, another civic actor said:

> I think my primary concern is that it potentially drives more of what's already happening that's bad. This is just exacerbating the offline racial profiling or the offline victimisation of minorities. They are usually over-policed and under-protected. (social justice actors)

It is worth noting that situating data-driven policing within the historic and contemporary mandate of the police foregrounds that it is not the who and the what of policing that is changing, but rather, it offers another tool in a long history of oppression, where the political and operational choices that determine which communities and crime categories are subjected to increased surveillance are seen as heavily racialised. As another civic actor puts it:

> The scanners that can scan mobile phones or some of the hardware in itself seems neutral, because you could use it on anyone, but we know that it will be used in the same way that police already police. So with a disproportionate focus on particular communities and places. (racial justice advocate)

Here, the use of mobile fingerprint scanners in the UK is referenced to demonstrate that seemingly neutral technologies become tools for oppression within a racialised policing context, as they will primarily be deployed against certain communities. The act of centring race within the discussion about data-driven policing reveals a tension into how civic actors approach the root cause of discrimination: the technology that reinforces and intensifies who is subjected to the gaze of the police or their mandate that becomes enacted through data systems.

Those who work closely with communities affected by police violence and social and racial justice actors who centre race at the centre of their analysis further situate the injustice of discrimination into the multitude of ways in which it impacts the livelihood and life changes of targeted communities. A community actor observed that the societal belief that data offers a ground truth about criminal behaviour inscribes meaning and justification into state actions, well beyond the police.

> What we see is almost a sanitising of police intelligence. I think that's what's taking place here. When the police turn up to court or what CPS [child protective services] rely on is, 'this is what the technology has told us'. (community actor)

The sanitising of police intelligence materialises through the process in which a person becomes the subject of the algorithmic gaze, which in itself warrants being subjected to increased police attention, and in turn, will justify future actions by the state. Referencing Shirley Anne Tate's (2016) article, 'I can't quite put my finger on it: Racism's touch', the community actor describes the more invisible and structural harms of data-driven policing as an optimisation of a policing mandate that has historically racialised crime.

> We almost become devoid of a language of trying to understand those harms that are invisible. (community actor)

> A question that I've been struggling with since doing that piece of work is the extent to which practitioners themselves can be insulated from the dominant discourses that are bounded in society. (community actor)

These observations foreground the struggle surrounding the injustice of discrimination. The normalised racialised discourse around crime emerges from racially inscribed social norms and values that structure European societies, which in turn are internalised by individuals, communities, and organisations. Stereotypes, such as 'the gang member', 'the terrorist', 'the gipsy', and 'the poor', become 'part of a collective memory within those institutions and organisations' (community actor). From this perspective, data-driven policing is part of a catch-22, where the organisation, its officers, and data-driven policing continuously reinforce the collective consciousness from which they emerge, further inscribing normative notions of who should be classified as deserving or undeserving citizens. There are immediate consequences of this catch-22.

> It's becoming a matter of social control, and police brutality, police violence, becomes also a normalised manifestation of institutional racism, and nobody challenges that. What we see is that racial profiling does not stop on a document check. When people resist, it leads to real police brutality. (anti-discrimination advocate)

Here, the affordance of technologies is to further neutralise a certain organisational practice, in this case, identity checks, and resistance against this form of oppression is perceived as disobedient behaviour towards the legitimate authority, which in itself can condone a certain level of force and police violence. Even when the encounters do not end in police violence, there are other material effects of police stops. 'It's not only a control, it's not only two minutes lost in your day, it could lead to death, but it could also lead to many other ways of dehumanising people' (Belgium human rights actor). Listening to a broad range of voices shows that there is not one but rather a stratified

understanding of discrimination; some foreground it as the result of data-driven policing, while others centre it as a key organising principle of police oppression. It is those who work alongside and with targeted communities who bring to life what it means to enact discriminatory practices through surveillance technologies. It determines how an individual is perceived by the state, society, and their community, and as such, it impacts their sense of place-making, belonging, and worth in society.

7.1.2 The criminal justice trap

The second site of concern relates to pre-emptive state intervention and the long-term impacts on those communities that are already disproportionately subjected to its gaze. What civic actors refer to as the *criminal justice trap*, a term that aims to challenge the positive connotation attributed to state interventions, to argue that pre-emption does not deter but rather is aimed at controlling and excluding some individuals and communities from society and becomes a vehicle into the larger criminal justice system.

Police intervention, according to civic actors who work with impacted communities, creates a downwards spiral. ‘The criminal justice system is becoming even more of a trap’ (privacy advocate) and ‘to be criminalised or to be given, therefore, that early conviction, caution, whatever it is, almost facilitates the next stop and search. It becomes a vehicle into a criminal justice system’ (community actor). Similarly, a human rights advocate working in Belgium explains:

> What I’m very concerned about is that the state will make increased attempts to draw young people into the criminal justice system at an earlier age under the guise of prevention and deradicalisation. We find ourselves in a situation where children, in a socially vulnerable situation, have many opportunities taken away from them.

Minors and young adults do not randomly become the subject of state scrutiny; these actions follow active policy choices that are aimed at identifying and controlling specific behaviours. Here, early interaction with the police, early conviction, and being labelled as at-risk will increase state scrutiny and limit their life chances. To fully grasp its negative impact, I argue, we need to situate early intervention with the multitude of subtle and often invisible harms that emerge from it.

> What struck me was how profound those encounters with the police are for young people. There’s a significant intrusion that impacts upon young peoples’ sense of

> belonging. It almost forces them to ask questions around what they've done wrong. (community actor)

The internalisation of not belonging is the real but often invisible physical, social, and emotional harm that emerges from actions taken on data-driven policing outcomes. Police interventions are seen to reinforce this idea of the other, being part of a community that does not conform to the imagined notion of what society should look like, and by merely existing, they must be doing something wrong, which can make young adults question their sense of belonging and self-worth.

> When individuals are suspicious, or they are different, or they do things differently to what is perceived to be the White imagined norm, then, yes, they are subject or liable for intervention, positive interventions to support those individuals but those interventions can be increasingly coercive. (community actor)

The gaze of pre-emptive policing originates from somewhere, from the dominant norms and values that govern society, which is positioned as the White imagined norm that codes certain behaviour and bodies as different, which in turn justifies subjecting them to certain interventions.

The neutralising effect of data-driven policing, described in the section on discrimination, is seen as a tool that further intensifies the influence of the police in a racialised and impoverished person's life. Small events that are registered in police databases can justify increased state scrutiny on both the individual who 'messed up' or 'doesn't conform' and their environment.

> Punishing partnerships are encroachment into the family as a way of leverage control over the individual and a mother, but secondly, it's indicative of the encroachment of the state into Black and Brown lives and working-class lives. We literally can develop a range of tools and strategies as a way of trying to build compliance and conformity. (community actor)

Crucially, when connecting data-driven policing to the racialisation of crime, it shifts the narrative away from injustice embedded in data to how our governance structures are designed to control and exclude certain communities from society. The value judgement of who is considered a good or immoral citizen as such produces a pre-emptive crime approach that is designed to change and conform the behaviour of racialised and impoverished individuals through intervening in their lives and environment. It is devised to be coercive and spill over to the entire family.

> Those interventions are increasingly contingent on a number of social goods, and in essence, that's where the tension is. We're opening up the family to surveillance more and more and that's difficult. (community actor)

These increased surveillance and coercive practices on the entire family make social goods contingent on cooperation with the pre-emptive approaches. It's insignificant if these are actual conditionalities as articulated by Henman (2011), where social welfare is increasingly becoming dependent on the cooperation of an individual or family with other state practices or if these are perceived conditionalities, it creates a culture of compliance and conformity. Repression within the criminal justice trap, thus, relates less to physical force and more to control through coercion, where resistance is met with increased attention from the state and compliance is rewarded with increased access to social services.

7.1.3 Data protection issues

The third site of concern relates to data protection issues of data-driven policing. Civic actors articulate stratified privacy concerns: the heightened interest in accessing and storing increased volumes of data for a range of policing purposes is perceived to be problematic in itself and is connected to an increasingly asymmetric power relation between the police and the public, and the broader chilling effects of surveillance. I will first explore the data protection issues and then situate these in relation to their broader effect.

There is a sense that embedding new technological capabilities within policing heightens their interest in accessing increased volumes of data from their own data collection practices, offered by other public authorities under the guise of multi-agency partnerships, and collected from public and commercial data infrastructures such as ANPR systems and social media platforms.

> In terms of the actual technologies themselves, my first concern is always the data. The amount of data that authorities are able to access, or draw on, and the types of data which are often very spurious with regards to their aims. (privacy expert)

> We are actually seeing a very strong evolution towards more and more data, lots of different kinds of databases, trying to cross-reference as much as possible, trying to get as much information out of them as possible. (human rights expert)

Data-driven policing is perceived to allow the police to expand the nature and scope of evidence collection and surveillance beyond that of their own activities, which further infringes on the individual right to privacy. The other data protection issue is that of data retention for purposes other than investigation and intelligence.

> Regardless of whether you meet the risk profile or not data is kept for an extended period of time and further processed in a separate database for police learning purposes. (human rights advocate)

Data collection and retention serves multiple organisational interests, one of which is for the purpose of organisational learning. This raises the question of how this expanding nature and scope of evidence collection and surveillance is justified under Europe's data protection laws. Scholars have pointed to the loopholes and broadly defined exceptions for police data-processing purposes within the LED (Kindt, 2020), while civic actors specifically point to the lack of compliance. 'The police are very bad at following the law, and that is especially true of the data protection regulations governing the use of data by police' (Dutch privacy advocate). In chapter 4, I observed how an internal police audit in the Netherlands revealed that none of their critical data infrastructures complied with internal police policy nor with the law. This runs the risk that data is processed for policing purposes that are beyond what the data was originally collected for.

Not complying with data protection regulation is perceived to lead to a deteriorating information position of individuals vis-à-vis the police, which underscores the scholarly notion that datafication and opaque data practice are creating an information asymmetry between the data subject, the data collector, and the data processor (Andrejevic, 2014). Data collection about a person from everyday data infrastructure for the purpose of policing runs a risk:

> You almost have a scary scenario where you don't really know what the police know about you, and you don't really know the evidence that's going to be used against you, and even up until that point, you've got no way of actually being heard in policing practices to say, actually this data you have on me is not accurate, or you've drawn conclusions or assumptions from my data which is unfair, inaccurate, irrelevant, or something else. (legal practitioner)

In the context of policing, the concern about the deteriorating information position of individuals points towards an injustice in which the state can almost invisibly encroach upon the private life of

their subjects. This curtails individual agency to control who knows what about them. In this context, a privacy expert points us towards the notion of intent.

> It is an infringement on our autonomy, the moment surveillance is used and prevention takes place, then in some way, people's behaviour is changed; otherwise, it would not be prevention. The police steers towards expected outcomes. Irrespective of your political opinion on this, that is a lot of power in the hands of an authority to be left unchecked. (privacy expert)

Data-driven policing as such is an extension of the state desire to control society, where the logic of prevention connects the collection of large volumes of data to police interest in pre-emptively interfering in and influencing the lives of some people. This stratified approach to privacy connects data protection issues to broader questions of power, specifically who gets to determine what possible data futures look like.

7.1.4 Governance concerns

The fourth site of conflict is primarily foregrounded by those civic actors who work for digital rights and human rights organisations. They raise a number of governance concerns: the practice of pilots, the lack of transparency, and the broader state's belief in technology. In chapter 6, pilots emerged as a practice that allows police to speed up the introduction of newer technologies within the organisation. For civic actors, pilots are a political frame that allows police to gain new powers without proper due diligence, public scrutiny, and oversight.

> Very worrying trend. There is a culture that loosely assumes that, because this is a pilot, experiment, or testing ground, you don't have to be so strict with human rights standards and data protection laws. (human rights advocate)

Another civic actor observed:

> In a context designed by the police themselves, these pilots are often very unclear – how long it will last, what the exact research question is, and when it will be transformed from a pilot into a situation that will last longer. By then, it has already become so normalised that there will no longer be a debate about whether this is actually desirable and whether there should not be a legal basis for this. (privacy advocate)

Pilots as a frame are seen to create a legitimising dynamic around data-driven policing. In the first place, the frame legitimises certain technological experiments by reducing the perceived legal requirements police need to adhere to. Then, the operationalisation is designed by the police for the police, which is seen to create opaque conditions under which these experiments take place. When these pilots become a self-proclaimed success, the fact that the police have already been deploying these technologies legitimises future practices.

Civic actors are critical about pilots as a practice to open up space to 'get things moving' (privacy advocate), referring to the affordance of pilots as an internal operational vehicle that I discussed in chapter 6. Their critique relates to the perceived legal exceptions pilots are seen to create and to what end these experiments are deployed. 'What I found most amazing is that it is really seen as an experiment for the police to learn from, instead of really thinking from the start about what this means for citizens' (human rights advocate). This observation points to friction between the presumed social values on which the police can legitimatise the turn to data-driven policing and possible infringement on the rights of citizens, the public good, the security and safety of people, and the reality in which the police see data-driven policing as an organisational optimisation tool. Finally, civic actors situate the pilot frame within its broader context. 'This is a broader trend that we see more often when it comes to large government data-processing projects' (human rights actor), where this emerging practice is not limited to the police. 'At the moment, there's nothing stopping these trials from going ahead' (privacy advocate). She hints at an enabling environment that creates a certain dynamic in which there is 'no one really in the driving seat, but there is a lot of petrol in the tank' (privacy advocate). These broader interests in the use of data systems create barriers for action, as civic actors have to challenge its use and the broader interest of the state in data-driven decision-making.

Another key governance concern is that of police transparency. Where debates in critical data studies often refer to the transparency of algorithms (Ananny and Crawford, 2018), civic actors raised this issue in relation to democratic processes and procedures. A British privacy advocate argued that 'the kind of data and privacy concerns is the processing [for] which there is no transparency'. A Dutch privacy advocate notes that 'there is an immense lack of transparency about what the police actually do. For example, with these pilots, they just go ahead and do it, and there is actually very little advance notice and very little debate surrounding it'. A Belgium human rights advocate argues that

> our biggest problem is a lack of ex ante transparency. There has not been a single parliamentary debate about the iPolice idea, nothing, the government has announced it, it has launched a very small press release, and the iPolice system is being set up and there is no debate, no public debate about it.

All the observations draw attention to the notion that the police implement data-driven technologies without or with very little public deliberation or advanced notice. The lack of proactive communication puts the parliament, civic actors, and the public at large in a disadvantaged position, as their response can only be formulated once a specific tool becomes practice or when the police put out a press statement, as was the case with facial-recognition trial in Zaventum. There are also complaints about the lack of ex post transparency,

> I submitted an FOI [Freedom of Information] request. Now it is a year later, and all I have received is a master's thesis written by a student. The police themselves say that they have no other documents. (Dutch privacy advocate)

> We have a very complicated state structure, and very often this is abused to make things very complex so that no one knows who is responsible for what; if you are disadvantaged, you don't really know where to turn. (Belgium human rights advocate)

There is a distinct feeling that police close ranks and actively prevent information sharing with the public, and in the context of Belgium, make use of the complex governance structure to obscure who is responsible for certain practices. This injustice of actively creating an information asymmetry between those in power and their watchdogs or those affected relates to the concern of unchecked police power. 'If you don't know exactly what is happening, you can't ask concrete questions about it, and you don't even know what problem to raise'. The human rights actor articulates how the lack of transparency leaves civic actors and the public in the dark and thus limits public debate as an accountability instrument. This is not a new policing challenge but rather points to the notion that the emergence of data-driven policing is also subjected to the historic information asymmetry between the police and the public.

Another key governance concern is how police influence and are influenced by their political environment. A Belgium human rights expert argues 'as in many areas in the society, there is some kind of a technological solution sanctioned, and the technology is there, and we think it's going to solve the problems so we use it'. He points to the broader political rationale that privileges the use

of technology within public services, which in turn normalises police use of technology. Another civic actor observed:

> The parliament says in one sentence: yes, we think privacy is important, but we must be tough on it [crime]. Well, then you burden the implementing authorities and the legislators with an impossible task. (Dutch human rights actor)

Civic actors observe that the political discussion on the conditions under which these tools can be used are top-level and offer conflicting messages, which leaves a lot of room for an implementing body, such as the police, to interpret how to balance the use of these technologies against possible privacy infringements.

Again, the lived impact emerges when racial and social justice actors situate data-driven policing as a technology of governance in the broader political environment of the contemporary welfare state.

> If they could prove to me that they suddenly had all these resources for therapeutic support that was going to those people on the list, that would be one thing, but this country has been decimated by austerity. So this notion that it's safeguarding for those people, I think, is really fundamentally dangerous because it's selling hard enforcement with this idea that we're helping them and it isn't helping them. (racial justice advocate)

Similarly, a social justice advocate argues:

> I'm in any way suspicious of increased budgets for policing, particularly in an era of austerity and increased criminalisation of racialised people and working-class people. So in any case, there's scepticism from many of us who work in the racial justice field about increasing policing budgets. (social justice advocate)

In chapter 4, I discussed how, in the context of the UK, the police and other public services have been confronted with steep austerity measures. Experts have argued that this has exposed police to fill the void that is left by other public institutions. Think of responding to the increased levels of homelessness and individuals with mental problems that are on the streets, which requires them to step in and deliver services they are not trained for. This reality makes it difficult for civic actors to believe that the care and control approach will actively invest more resources in the care components of prevention, but rather mask hard law enforcement under the guise of prevention and support. When funding becomes available for the police to pilot data-driven policing technologies

in an effort to do more with less, without also investing in other public services, these support interventions thus exist merely as a political frame.

7.1.5 Chilling effect

The final concern I will discuss in this chapter is that of chilling effects. While only a few civic actors use the word *chilling effect* to describe the impact of data-driven policing on society, it is clearly a concern. A legal scholar draws our attention to two distinct types of chilling effects: ‘They don’t talk about it, the chilling effect. You go to a demonstration, you go somewhere, and you know that you will be recorded and put in a database for ages’. He talks about the chilling effect in relation to the rights to protest and assembly that can prevent a person from becoming or continuing to be an active political member of society. He continues by referencing Virginia Eubanks work on *Automating Inequality* (2017) to articulate that, when data from other public authorities becomes part of policing interventions, it creates barriers for people to access social services and social rights:

> If you can afford to put your child into private care, then you’re going to do that to escape the eyes of the law; essentially, escape the over-policing. You have a really fundamental question there about your right to social security, your right to an adequate standard of living, if you are forced away from these other types of government support or other types of state support mechanisms because of the policing implications.

The concern is that this practice will entrench class structures in society; citizens with enough resources fall outside of the realm of state scrutiny, and those who rely on state services have the choice between the immediate needs and the longer-term repercussions of depending on it. Similar concerns have been raised in meetings organised by civic actors about data-driven risk-scoring programmes. Stories were shared about cases in which families made calls to the emergency services to de-escalate domestic violence situations that resulted in children having additional data points tied to their name in police databases, flagging them as being at risk and subjecting them to pre-emptive intervention on behalf of the state. These experiences resulted in an overall decrease of trust in the state and deterred families from trying to get protection and assistance in succeeding emergencies, leaving them exposed to insecure situations. Therefore, the notion of chilling effect can be connected to what some suggest is a flawed logic of data-driven policing:

> Predictive analytics shows a fundamental misunderstanding of the complexity of peoples’ lives, family lives, upbringing, psychology, early experiences. This is not a

> case of really listening to peoples' experiences and people's needs, it's a case of broadly stereotyping and profiling people in many cases. (privacy advocate)

This reflects the top-down approach to data-driven risk scoring that allows police to sort individuals into categories of who fits within a normative construction of criminal or at-risk individual, rather than starting the intervention from a community perspective, listening to their needs and perspective on crime. A human rights advocate explains his concerns relating to social sorting: 'suddenly, citizenship gets turned on its head. You're no longer like a free citizen and innocent until proven guilty, but you're a suspect walking around, just waiting to be found guilty of something', shifting the continuous gaze of the state upon the individual from a place of distrust, which erodes the foundations of democracy.

I will conclude the first part of my chapter on civic actors by summarising my key observations. I have explored different themes that emerged from the concerns raised by civic actors in relation to data-driven policing practices: discrimination, criminal justice trap, data protection, governance, and chilling effects. By engaging with these topics, I argue that, while all civic actors engage with questions of power when critiquing the police turn to data, there is not one single understanding of what is at stake. Rather, my findings highlight how there are a range of concerns, and civic actors inscribe different meanings into each of these issues. Take, for example, the meanings given to the issue of discrimination; some actors relate to it as an outcome of technology, while others position it in the historic and ongoing racialisation of crime. Similarly, the impact of the increased volume of data made available to police is connected by some to data protection infringement and how this practice increases the power asymmetry in the police–citizen nexus, while others relate it to chilling effects where the desire to escape the gazing eye of the state results in an infringement on a broader set of rights. Also, the notion that data-driven policing create a continuous gaze that puts citizenship on its head presumes that targeted communities had a different experience prior to its introduction, which in itself conflicts with the articulation of harms that emerge from the criminal justice trap. While the injustice claims at times overlap, conflict, or operate alongside each other, bringing together a wide range of civic voices shows that their understanding of the nature of data-driven policing and its relationship to police power is dependent on their own histories and experiences. In the rest of this chapter, I will refer to these distinct entry points as the privacy, human rights, and race lenses. Below, I will use these lenses to explore the hierarchies and tensions between them to gain an understanding of how they at times align and conflict, whose voices and concerns count and get privileged in the discussion on police power, and how this relates to questions of resistance or reform.

7.2 The politics of injustice

This chapter has contributed to my research question on the relationship between datafication and police power by explaining civic concerns on the introduction of data-driven policing. The different injustice claims foregrounded by a range of civic actors show that the use of data by the police does not create one but rather multiple sites of struggle. In chapter 2, I drew on social scientists to argue that police power is relational, in which it is a proposition made by a power holder that is responded to by different publics, which in some cases requires a response from the power holder. As such, my research into police power aims to account for its relational dimension by exploring how the multiple sites of struggle affect which voices get privileged in the understanding of just and unjust policing. Therefore, in this next section, I will explore the dynamics around the at times overlapping and conflicting privacy, human rights, and race lenseses. First, I will explore how the changing composition of civic actors working on police power is perceived to create an invisible hierarchy into whose voices count and which injustices are privileged in the debates on data-driven policing. This is what I will call *the politics of injustice*. Second, I will offer insights into the external dynamics that are perceived to create a need for social and racial justice actors, and to some extent human rights actors, to engage with questions of technology and police power. Third, I will explain how the use of data by the police influences civic practices of reform and resistance to argue that the politics of injustice raise competing injustice claims that police can choose to respond to. I conclude that the introduction of data-driven policing impacts those subjected to its gaze and the civic space by privileging those voices who centre reformist technology agendas over more abolitionist forms of agency that challenge the racialisation of crime and police oppression of Black and Brown communities in the debates on just and unjust policing.

7.2.1 Emerging hierarchies and tensions

To gain insight into the alignments and conflicts between the different lenses, I will start this second part of my civic actors chapter by listening to the experiences of those closest to the people who are systematically oppressed and controlled by police interventions: the racial and social justice advocates who have a shared identity, work with or represent targeted communities, and enter into the discussion on data-driven policing primarily through the race lens. Their experiences with the broader human rights field, and digital rights advocates and technologists more specifically, foreground a dynamic in which those voices who experience and have traditionally worked on discrimination and policing are pushed towards the margins of the debate, and technology is centred

in the discussion on just and unjust policing. Thus, I will argue that centring data in the understanding of police power creates invisible hierarchies between civic responses, what I will refer to as *the politics of injustice*.

Race as an entry point into the discussion on data-driven policing allows civic actors to explore how policing priorities emerge from a complex set of social norms that determine which individuals, communities, and behaviours do not conform to a dominant idea of what a European society looks like. A community actor observes:

> So my first question has always been: who is this designed to exclude? The tech is designed to exclude those bodies, those individuals who don't look, who do not conform to that normative White imagined society or those individuals who don't conform to notions of class and expectations of those individuals around those spaces. (community actor)

Technology as the manifestation of broader social norms that exclude racialised and impoverished communities demands an inquiry into how these normative notions of what society should look like affect civic action. In the previous section, racial justice actors explained that it is difficult for police and public servants to isolate themselves from the tropes and stereotypes that are deeply embedded within the process of othering and exclusion. Below, I will explore what these dominant social norms mean for the interaction within and between civil society.

A racial justice advocate reflects on her first encounter with the digital rights community, in which she and one other person were the only people of colour.

> I was really shocked. You would think that all these White people in the room had just discovered racism. This one guy was like, this opens all sorts of doors for us to discuss racism with the police, and I was like, what do you think communities have been doing for forty years? The lack of understanding of peoples' lived experience but also the lack of historical understanding of race and racism and even policing was jaw-dropping. It was as if these tech people have suddenly jumped into policing and race without any background or context.

This experience highlights a dynamic in which actors who are part of the privacy and technology discussion are interested in engaging with a new issue, the discriminatory impact of technology in the context of policing, and fail to recognise the historic relationship between racism and policing.

The act of centring technology in the issue of discrimination and policing without actively trying to understand and work with those who have been working on dismantling the structures that racialise the construction of crime thus creates invisible hierarchies between what knowledge counts and reduces the discussion about power and oppression to abstract notions of bias. She continues to argue that

> there also needs to be a lot of training around race and class history for people. There were all these abstract conversations, but I don't think there was an understanding of what data harm actually means to someone impacted. […] So that was my first experience, and you can see I've been damaged and not wanting to go back.

Again, she points to a lack of knowledge by those actors who centre technology as an object of study on the histories of policing and race and its immediate impacts on Black and Brown communities in Europe as well as its impact on social and racial justice actors. This ignorance inadvertently makes them feel like outsiders in their own struggle and unjustly burdens them with the emotional labour of educating other civic actors on racism and policing. Similarly, social and racial justice actors who work with marginalised communities said:

> You actually have to include them from a strategic level at the start. Your work has to be shaped by it; otherwise, it doesn't work, and it's tokenistic, and it's extractive and all of these things. (social justice advocate)

The power dynamics within and between civic actors speaking on behalf of communities or not including them from the start constrains them 'to have agency in shaping solutions that centre their realities, approaches and knowledge' (Equinox, 2021, p. 17). This has broader implications for how we come to understand justice in a datafied society, as it raises the question of who has the opportunity and the power to shape how injustices and solutions are defined.

The effect of centring technology in discussions on police power inadvertently displaces racial and social justice concerns on the racialisation of crime and police oppression, and it is not a conscious act but rather the result of civic actors' positionality and the internalisation of dominant norms and values. A community actor reflects:

> it's not a tough one. I was going to say they don't get it, but that's not fair. If I go back to the language of 'Racism's Touch', the encroachment of the state into Black lives, that sense of being perceived as different and outside, White digital rights activists, White

> surveillance experts often don't get it and won't get it, and these spaces are predominated by White folks.

Here, he points to the issue that, if one never fundamentally engages with questions of power, race, and oppression and if one has never lived through these experiences, it is difficult to understand the physical, emotional, and mental harms that emerge from the collective punishment of racialised communities. He continues to argue that

> the common-sense view that terrorism is perpetrated by young Muslim people because they're pissed off at Western societies. It's difficult to disrupt those tropes. I can understand them also accepting those dominant discourses and those narratives. It's really difficult to shift away from them.

He foregrounds how civic actors are themselves not insulated from the racialised construction of crime, as they are part of the dominant culture and collective consciousness that shapes society. Shifting away from this requires actors to actively unpack and oppose these tropes and stereotypes. A race lens sheds light on the impact of agents positionality, as individuals or groups can be highly reflective in one area but uncritically reproduce attitudes and behaviours in other areas, challenging the norms of surveillance and state power while at the same time reproducing stereotypes and racial injustices. He continues to argue that

> I'm thinking you haven't asked about the human rights. I've never fucking been afforded human rights in these spaces. I think it was Isaac [scholar Walter Isaac], wasn't it, who speaks about the imagined White community. The tech is designed for them.

The notion that the human rights ideology assumes that everyone is afforded human rights in Europe, which in itself reproduces racialised discourse, as it advocates for rights and changes to technology that exclude parts of society. 'So I work on the basis that digital rights organisations or human rights organisations are also not there to support Black and Brown communities'. Thus, unless civic actors actively engage with the race lens, they cannot properly support the communities that are affected by data-driven policing technologies, as they do not address the root causes of discrimination and exclusion. These different accounts speak to the invisible hierarchies social and racial justice advocates are experiencing when engaging with the broader civic field. The introduction of data-driven policing creates a politics of injustice in which discussions on technology and human rights are seen to displace crucial discussions on race and police oppression. In addition, civic actors that lack a critical understanding of racialised patterns of oppression and

control have displaced discussion on exclusionary police practices and have made racial justice advocates feel like outsiders in their own fight.

7.3.2 The need to engage with data-driven policing

The introduction of data-driven policing is creating multiple sites of struggle, which create competing narratives on how we come to understand what just and unjust policing looks like. These civic dynamics, I will show, do not operate in isolation from their broader context; rather, they get reinforced by a number of external forces. In this next section, I will discuss the external social and institutional structures that are believed to influence which injustice claims get centred in the discussion on data and police power.

There is a sense that not engaging with questions of data and technology is perceived to close opportunities for finances, action, and public interest. Civic actors who work on these topics from a critical race lens argue, 'you can't avoid it either', and 'I can't afford to tell the digital rights folks to do one. We literally need them'. These observations point to a dynamic in which police investment in data-driven policing creates a necessity for those who are part of and working with targeted communities to engage with those who are knowledgeable about data. Here, I note that the inverse dynamic, the necessity to work with social and racial justice actors, was not explicitly articulated in the interviews with civic actors who approach data-driven policing from a privacy or human rights lens. In addition, social and racial justice actors feel that there are financial incentives to engage with the police turn to data. A social justice advocate observed:

> Yet politically, if you look at the money and you look at what is more likely to be funded, if an angle doesn't have a digital element, it's very hard to get. This isn't a sexy issue right now, and the topic of digitalisation is sometimes much easier to get attention into and also get funding and resources for than if you look at just the base core issue in itself. (social justice advocate)

Thus, the civic space reflects broader social and institutional structures, in which hierarchies between the different approaches to injustices are reinforced by funding mechanisms that support those activities and organisations that work on the 'hot' topic of technology.

The privileging of data protection and privacy arguments over other human rights violations are also observed by other civic actors. A Dutch human rights actor argues that 'personally, I think it's a shame that very quickly the discussions become about privacy, because I don't think that's the

biggest problem we face. I think automated inequality is a bigger problem'. Here, she indicates that there is a broad range of human rights issues at play when the police turn to data-driven policing, and focusing on privacy runs the risk of overlooking or overshadowing those issues. This resonates with the observation of a legal practitioner:

> A lot of the litigation has been narrowly tailored towards privacy and data protection arguments, and I don't think that's necessarily a bad thing, but it misses other arguments, like other impacts of this technology on other human rights […] Courts will often go for the human rights arguments that are the easiest to reach a decision on, and they'll even say in their judgements because we've reached a decision on privacy, we're not going to consider the discrimination arguments, or we're not going to consider the right to assembly arguments.

Here, he notes two distinct dynamics that privilege a data protection argument over a discrimination argument within strategic litigation as an avenue for action. First, the known litigation cases against specific sociotechnical assemblages have focused on a narrow data protection argument, and even when cases include legal arguments that include both data protection and discrimination arguments, external structures like the courts favour ruling on the first over the latter. This indicates that, while a human rights lens allows for a broader analysis of power, there is an unspoken hierarchy in human rights, in which the judiciary and executive branch of the government will go for the arguments that allow for clear and uncontroversial ruling and sidestep any hairy issues.

Listening to the experiences of those engaging with data-driven policing through a race and human rights lens show the myriad of ways agency reflects the dominant social norms and values of society, which inadvertently create tensions and hierarchies in the field of civic action. It has to be recognised that not all the experiences of racial justice advocates engaging with the digital rights field on issues related to data-driven policing practices have been negative, especially those actors who say they are relatively new to these discussions:

> Always had very open conversations in very safe environments. Me being a person of colour, it's always difficult to ensure that we are not making too much controversy on something, that we want to engage with our lives, not disengage, but it was always very open. (racial justice advocate)

> There's a kind of magic between the digital world and the anti-racist world because it's very new. (anti-discrimination advocate)

In part, this positive sentiment can be attributed to the fact that the digital rights field is evolving: efforts have been made to increase cooperation with other human rights and social justice actors, to diversify the digital rights workforce, and to engage in discussion around decolonising the field (Digital Freedom Fund, 2021). Some digital rights actors are aware of the implicit power relations within civil society, between civil society and targeted communities, and between civil society and the state. They foreground a dynamic where

> you almost never see any Muslim organisations being invited, any NGOs that focus on anti-racism. Like you see us, and that's already a step forwards. There is also a huge imbalance between what type of organisations are there. (digital rights advocate)

The observation of whose voices are heard shows that policymakers in Brussels primarily consult digital rights groups in matters concerning new technology, regardless of whether it intersects with broader human rights or discrimination issues.

So far, this section has engaged with the external dynamics that create a need for civic actors to engage with questions related to technologies and the new barriers these external forces create. Now, I'll engage with external opportunities that are seen to emerge from the use of specific data-driven policing functions. The use of data-driven policing is seen as an opportunity to position its injustice as collective harm that will impact everyone's rights.

> Things like facial recognition, or predictive policing, I think we could get more support from the general public, because the impact is really on everyone and not only on terrorists, or would-be terrorists, or other people persecuted. But really facial recognition can impact everyone every day. (human rights advocate).

> Potentially, it would be everyone's privacy, which I think is important. (racial justice advocate)

In this sense, data-driven policing has the potential to affect the general public, who up till now, have had limited to no first-hand experience with the police, rather than only the historically racialised communities who have been at the receiving end of policing actions. The need to engage with the public debate around facial recognition relates to this notion of collective harm and the hope that this can create a shared experience that will bring about collective reflection on the legitimate use of technology in policing. Those actors who have historically approached issues around mass surveillance from the perspective of privacy are more sceptical about this notion of

collective harm. ‘In the Netherlands, the debate is very apolitical at times, and there is a deep-rooted trust in the police’ (digital rights advocate). She continues to argue that

> in fact, everything the police stand for is seen as a great public good and this is then weighed against privacy, which is seen as a kind of individual right. And that you, as an individual, just have to make that small sacrifice for the greater good of society.

These experiences show the difficulty of positioning privacy as collective harm, as it runs the risk of positioning privacy against the dominant security discourse. What makes it even more complicated is that this debate is skewed in favour of the police, as people have an intrinsic level of trust in the police as the legitimate state actor that holds the monopoly on domestic violence. In addition, the notion of collective harm presumes that, if the police engage in mass surveillance activities, everyone will equally be confronted by it. In the context of the Sensing project of the Dutch police, a human rights actor challenges this assumption:

> So sensors are placed in one part of the city and everyone who drives around there is scanned, regardless of whether you meet the risk profile or not, and everyone’s data is further processed in a separate database and stored for a month for the learning purposes of the police. If you meet the risk profile, your data will go through another route, where it enters the operational database and a report is made to the police officer, and if the police officer decides to follow up on the report, the data is transferred to the regular police database. (human rights advocate)

This example shows how mass surveillance might record data from everyone within a specific area, the choice that informed the intervention, the actions that are based on its output, and that the subsequent processing of this data is still heavily racialised, and as such, the possible harms are again borne by a small subset of people that meet the risk profile.

7.2.3 Practice of ‘reform or resistance’

As outlined above, civic actors express a multitude of concerns in relation to data-driven policing, and while all these concerns speak to the notion of police power, there is not a single notion of what is at stake that would have allowed civic actors to outline a set of universal principles. Rather, I have shown that data-driven policing is shifting the composition of civic actors who engage with questions of police power, which is creating invisible hierarchies between them. These developments do not materialise in isolation; rather, they reflect broader social and institutional

structures that privilege actions that address privacy concerns over other injustice claims about data and police power. In this final section on the politics of injustice, I will look at the actions repertoire of civic actors, which are organised along the lines of practices of reform or resistance, where the first is aimed at creating conditions to make things less bad and the latter more fundamentally challenges police power. I conclude that these practices create competing injustice claims the police can choose to respond to.

To highlight the distinction between reform and resistance, I want to put forward an observation made by a social justice advocate.

> I think another challenge is the willingness of people, of the movement in itself to say no, to just in general exercise refusal of certain practices or technologies rather than looking for reform. It's basically being like, a company's going to develop it, so it has to happen and therefore we just have to deal with it and find safeguards to make it better, rather than actually the much more empowering approach. (social justice advocate)

The perceived inevitability of data-driven policing also informs civic actors' action. There are two distinct approaches to action, including a more reformist agenda, in which solutions are directed at making existing and new power relations less harmful. Agency as resistance, also referred to by racial and social justice actors as an abolitionist approach (Benjamin, 2019; James, 2005), challenges the inevitability of technology and police power by refusing the dominant frame and trying to shift the conversation towards other narratives. While both approaches can be conceptualised in relation to empowerment, as it allows actors to influence the power structures around them, the social justice advocate hints at a difference in which reform is seen as actions that are taken within the parameters set by those in power rather than fundamentally challenging and changing these structures. She continues to argue that

> when you look at it from the perspective of marginalisation as opposed to privacy, you will see that many of the concrete demands won't solve the issue that you are trying to solve. If your aim is that you don't want people to be treated badly by the police when they're just living their everyday lives, then you have a different end goal and certain things will just not get you to that end goal.

In this reflection, she layers the act of resistance and reform with the understanding that civic actors each have a different articulation of the loci of struggle. Some see the emergence of data-driven policing and its privacy infringements as the problem, while others view it as another tool in larger

structures of police oppression and the racialisation of crime that materialises in concrete harms to some communities. This suggests that the concepts of reform and resistance are intrinsically tied to civic actors' problem analyses and beliefs in the legitimacy of the current power holders and structures in society.

Resistance against data-driven policing from a race lens as such is twofold: saying no to its use and challenging the loci of struggle that centres technology and not racialisation of crime in dominating discourse on it. From this perspective, there is a lot of pushback against technological reform, a term I use to describe the actions that approach data-driven policing as an isolated technical system and direct their efforts at minimising bias in oppressive systems. A community actor reflects:

> The notion of technological defensiveness, that the justifications for training the tech on particular groups in the communities is a way of improving the tech, is fascinating because you literally say, we're just trying to improve this shit. […] No. I'm not interested in you improving the accuracy of some of these tools, because these tools are being designed to be trained upon particular bodies, and they're invariably Black and Brown bodies.

When civic actors connect the historically determined racialisation of crime and its lived harms to that of data-driven policing, reformist practices are perceived to make these technologies seem more neutral and 'fair', but in effect, they optimise them to accurately target racialised communities. He continues to argue that, 'what we're doing is reaffirm who we see as a foci for attention and who has a presumption of criminality'. The discourse of training this technology to be more accurate on some voices, accents, and faces further entrenches the racialised discourse of crime and legitimises state investments into optimising oppressive technologies and their interference in the lives of some. Thus, the act of resistance that emerges from a race lens demands a different approach, one that unpacks and challenges underlying power relations from which these tools emerge and perpetuate, not the technology itself.

> Those abolitionist impulses speak to that the state and policing are focused and concentrating on particular crimes. If we were really interested in violent crime, surely we'd be responding to domestic abuse and sexual violence as experienced by women and children on a regular basis? (social justice advocate)

Resistance from a racial justice perspective is closely connected to the policing harms and insecurities that certain communities face and the need to shift the narratives from technology to the

political rationale that underpins policing and informs their priorities and actions. This resistance is a political act of agency that aims to unpack the racialised discourse of crime that has become embedded in the data-driven policing practices and in effect challenges the legitimacy of police as an institution.

If resistance relates to saying no and challenging the dominant narrative and social norms that inform police practices, how do actors from a race lens then engage with reformist practices? One approach is to align the social and racial justice agenda more closely to issues of technology and the digital rights agenda. 'This would maybe make the sphere of influence and the sphere of topics that come under the remit much broader and therefore more relevant to the root causes of the issues that we're trying to contest' (anti-discrimination actor). Reformist demands that challenge a data-driven policing system on its privacy and data protection infringements are seen to support the fight against more entrenched patterns of oppression and control that are driving these developments.

> Like it kind of doesn't focus on the parts that get people hot and bothered, and that kind of gets people's hackles up. But it's like, actually, this particular practice is just evidently contrary to GDPR. (anti-discrimination advocate)

> So we are trying to be more and more present to really ensure that the safeguards are protected, or to reduce the risk of violation of fundamental rights. (anti-discrimination advocate)

Experience informs racial justice groups that there is more political will to conform to data protection demands than to more politically controversial topics like anti-discrimination demands. In addition, advocating for any kind of safeguards to make data-driven policing less bad is perceived to alleviate the immediate suffering of some communities. This is not without conflict, as she continues to argue:

> That is where we struggle because this is a catch-22. You are there to think that you are providing something that is exceptional when really you are just able to protect some safeguards, and then you are trapped in a process that is just again normalising institutional racism. (anti-discrimination advocate)

Foregrounding that political agency directed at reform from a race lens creates sites for conflict, as these activities might mitigate some direct harms but by doing so further inscribe racialised notions of crime.

When we connect a privacy and human rights lens to the question of reform or resistance, I argue, different tensions emerge. At its core, civic actors who engage with data-driven policing from these two lenses find it challenging that they are operating in the unknown, 'a large part of my job is to warn of future scenarios' (human rights advocate), and 'we have no idea how these things are actually going to be used' (digital rights advocate). This suggests that technology as the loci of struggle, in part, requires civic actors to respond to eventualities of data-driven policing that might materialise in the future, which in turn creates unknowns about the actual harms. Therefore, actions are based on the experience that the introduction of data-driven policing functions will lead to human rights infringement and renew police desire to expand their surveillance practice. A privacy advocate situates it within the current legal regime that governs police:

> These human rights safeguards should start with a legitimacy check on the use of data-driven policing tools, as human rights infringements are allowed provided they are necessary, subsidiary, and proportionate, and in that necessity test, is very much should you do this at all. And actually, the answer is no, unless infringements are only allowed if they are actually necessary. (privacy advocate)

Here, she refers to the Necessary and Proportionate principles (Access et al., 2014) to challenge the presumption of the inevitability of data-driven policing. Thus, resistance, in this case, is directed at shifting the narrative from a focus on the conditions under which these practices can emerge to the legal frameworks that restrain its use, a shift from *how* to *no, unless*. A civic actor notes that a civic agenda of saying no runs up against organisational limitations:

> There's definitely a lack of political boldness, and obviously we're a registered charity and there are limits to what you can do politically. It's a very tiring agenda when every organisation is just making sure there are some new safeguards. At some point, I think it's just necessary to say enough is enough, but there don't seem to be many people who are willing to say that kind of thing. (human rights advocate)

As outlined above, questions of reform and resistance are dependent on the ability of professionalised civic organisations to say no. However, the emergence of facial recognition is seen to have changed the conditions under which these civic actors operate.

> We've signed a letter that EDRi [European Digital Rights] addressed to the European Commission. We didn't ask for a moratorium. What we asked for is a redline on many

> points. I'm more comfortable with this kind of demand, but I guess it's politically easier to ask for public debate and a moratorium. (human rights advocate)

> So practically everybody agrees that, on the political level, the call for the ban is the right one, because in many applications, there are no conditions under a moratorium that can mitigate or lead to a conclusion that would lead to something like, okay now it can be rolled out. (digital rights advocate)

Professionalised civic organisations have constraints in how they choose to engage with the debate. Individuals might prefer resistance, but organisationally advocating for softer measures will be easier and more acceptable. However, some technologies, like facial recognition, are seen to be so incompatible with human rights, where any reformist demands will offer police the opportunity to negotiate conditions under which future use is allowed, that resistance also becomes desirable from an organisational standpoint.

The limits of reform are also articulated in relation to it creating additional barriers for identifying harms and challenging future implementations. A digital rights advocate notes in relation to the bias debate that 'if it works badly, it's bad. But if it works really well and it's super efficient, it's awful. You're only perfecting the instrument of surveillance' (digital rights advocate). Here, he points to the fact that the important work to highlight the performance issues and error rates on recognising certain demographics in facial-recognition systems has led to reformist arguments that have resulted in the police investing in more accurate surveillance systems. This creates an additional challenge.

> So it's quite worrying to see when they adapt to the language and it makes them sound slick, and it forces us to be very careful about how we argue against the technology, that we're really putting the emphasis on the real harm and not on things that can be fixed and optimised. (privacy advocate)

Consequently, civic actors are wary of the ease at which demands for technological reform can be reappropriated to legitimise the conditions under which data-driven policing technologies can be used. This sentiment is echoed across the interviews. Civic actors who engage with questions of power through technology express a clear concern that calls for reform are creating the conditions in which data-driven policing will be legitimised.

At the start of this chapter, I showed how there is not one single understanding of what is at stake when it comes to the introduction of data-driven policing. Civic actors' distinct experiences and

belief systems, what I refer to as the privacy, human rights, and race lenses, foreground a range of injustices that emerge with the police turn to data. In this part of my chapter, I explored the alignment and conflicts between the privacy, human rights, and race lenses to understand the broader implications of data-driven policing on how we come to understand just and unjust policing. First, I showed that the introduction of data-driven technologies does not merely change the practice of policing, it also changes the composition of civic actors involved in challenging police power; it brings digital rights advocates and technologists into the discussion on discrimination and policing, and racial and social justice and human rights advocates into the discussion on technology. While in theory, the different lenses should align, as all challenge the enactment of police power through data systems, I find that it establishes invisible hierarchies between civic actors and injustice claims. The public debates on the harms that emerge from data-driven policing is seen to privilege knowledge on data over that of lived experience, racism, and policing. Second, I foregrounded that other stakeholders, such as civic donors and the courts, are not immune to the affordance of data and, through their actions, create conditions that privilege injustice claims that centre privacy over those of social and racial justice. Thus, the broader environment reinforces invisible hierarchies in the civic debate and makes it necessary for social and racial justice actors, and to some extent, human rights actors, to engage with the issue of data-driven policing. Third, I engage with the actions of civic actors, framed as the practice of reform and resistance, to understand how invisible hierarchies manifest in competing injustice claims police can choose to respond to. When approaching the injustices of data-driven policing, the race lens challenges the police as a legitimate authority and, by extension, their use of data, whereas the privacy and human rights lens challenges the legitimacy of the use of technology but does not engage with more fundamental questions around the police as a legitimate power holder. This has broader implications for how police respond to injustices foregrounded by civic actors.

7.5 Conclusion

To contribute to the understanding of the nature of data-driven policing and the relationship between datafication and police power, this chapter has highlighted the extent to which its emergence is changing the practice of civic actors and the injustices that materialise in public debate. Here, all civic actors are concerned with how these new practices provide the police with new powers without much public debate or oversight. The turn to data-driven policing is seen to normalise existing structures of oppression throughout the entire criminal justice system, create a multitude of infringements of citizens' fundamental human rights and right to privacy, and limit their ability of recourse and redress, as these practices are erecting barriers for access to justice,

social services, and information about government conduct. Indeed, when I listen to social groups, as suggested by Young (2011), a range of injustices related to data-driven policing emerge; however, the sum of these voices does not articulate one theory of data justice or even shared principles of justice, they rather point to a politics of injustice. Although it is often implied in the literature around social justice, when I connect the positionality of civic actors, what I refer to as the race, privacy, and human rights lenses, to our understanding of what is at stake, it offers insights into the emergence of invisible hierarchies in whose voices count in the debates on data and police power. Here, concerns that centre technology are seen to be privileged over those of the racialisation of crime and the legitimacy of police as a power holder. This dynamic manifests in distinct practices of reform and resistance, which, I argue, create competing injustice claims the police can respond to. As such, the introduction of data-driven policing has broader implications for how we come to understand police power and social justice.

8. Data-driven policing: A question of legitimacy

In this analysis chapter, I will draw on the insights discussed in my findings chapters to theorise about the relationship between datafication and police power. In my fieldwork, I explored the nature of data-driven policing in Europe as a practice and as a site of struggle within four contexts, Belgium, Brussels, the Netherlands, and the UK. Researching data as practice, I argued, allowed me to explain and expose how the emergence of data-driven policing is shaping approaches to crime, policing needs, and opportunities for action. I chose to look across and between European policing contexts as the practice is nascent, and there are only a limited number of projects and people working on it. This approach allowed me to identify broad organisational principles that structure police approaches to technology and offer insights into the social structures that (re)produce them. Drawing on social science debates that theorise on the dialogic nature of police power - it is enacted and responded to - I also engaged with an inquiry into civic debates on the introduction of data-driven policing to argue that its emergence is considered a site of struggle that has broader consequences for how we come to understand fair and just policing. The insights from my research allow me to contribute to the scholarly debates in two ways: to bring out the organisational rationale and social implications emerging from the introduction of data-driven policing and to make a broader theoretical argument about how power is justified and challenged in an increasingly datafied society.

To answer my research question on the relationship between datafication and police power, this chapter builds on three social sciences frameworks. First, Beetham's (1991b, 1991a) observation that the justification of police power is not so much about the belief in the police, but to the extent that the police and their actions can be legitimised along the dominant believes and values of a given society. Second, Bottoms and Tankebe's (2012) 'dialogic model of police legitimacy', and its refinements by Martin and Bradford (2021), which identifies three spheres in which police power is negotiated: power-holder legitimacy, audience legitimacy, and authorisation of external stakeholders (Bottoms and Tankebe, 2012, 2017; Martin and Bradford, 2021). Third, data justice as a relational lens to account for the struggle over who gets to define what just and unjust policing looks like (Dencik, Jansen and Metcalfe, 2018; Dencik et al., 2019). I will conclude my theoretical argument about the relationship between datafication and police power by putting forward the concept of *data legitimacy,* a relational vector to account for the distinct and stratified ways in which the turn to data is redrawing how we come to understand what just and unjust policing looks like. This will allow me to contribute to the field of media and surveillance studies by addressing some of the knowledge gaps identified in chapter 2. Where the dominant debates inform the understanding about the

shifting boundaries of who and what is seen as an authoritative voice in contemporary society and how datafication allows power holders to control present and future behaviours of people. However, I argued that these insights offer a limited view of the relationship between data and power, one that is less informative on its productive, normative, and relational nature.

In this chapter, I will first outline the key empirical findings on data-driven policing as a praxis and a site for struggle that have informed the conceptualisation of the concepts of *organisational optimisation logic* and the *politics of injustice*. Second, this chapter will build on the discussion on legitimacy and social justice to argue that power embodied and enacted through data systems has a normative, productive, and relational dimension. Here, I present the notion of *data legitimacy* to account for the multitude of ways in which data acts as a legitimising frame of what is perceived as 'just' policing. Furthermore, I will build on the concept of data justice as a relational lens to research audience legitimacy, one that accounts for both the injustices and for the politics of injustice that emerge from the introduction of data-driven policing. I conclude that these insights show that the datafication of society has broader implications for how we come to understand police power and social justice in contemporary society.

8.1 Data-driven policing: Its political

The ideological grounds of datafication and the rise of surveillance systems as a global phenomenon, I argue in chapter 2, have been well established. What scholarly debates from the field of media and surveillance studies have foregrounded is that data is becoming a significant feature of everyday life and is increasingly becoming a medium through which individuals and organisations, such as the state, the police, and civic actors, engage with the world. However, scholars have also foregrounded that there is still a lot unknown about the actual practice of data-driven policing (Brayne and Christin, 2021; Flensburg and Lomborg, 2021). This, I argue, makes the relationship between data and police power an important subject for empirical research. In this thesis, I understand the concept of power as the ability of an actor to manage or control the actions of others despite resistance, where power is socially constructed and productive, coercive, normative, and relational in nature. As such, my research approach engages with data-driven policing as a social process that is both an actual practice and a site for struggle to gain insights into the relationship between datafication and police power. Below, I will outline my key empirical findings from my mapping, risk-scoring, biometric recognition, and civic actors chapters. These insights will be structured along two main themes: the organisational optimisation logic of data-driven policing and the politics of injustice that create invisible hierarchies into which justice claims

count. To conclude, these two concepts foreground that data systems are shaped by and shape different dimensions of police power and, as such, have broader implications for how we come to understand the struggle over legitimacy and justice in the age of datafication.

8.1.1 Organisational optimisation logic

One of my key findings is that, from the perspective of the police, the turn to data is primarily framed as an organisational optimisation logic. Dominant scholarly debates often situate the political rationale from which the turn to data and algorithmic governance emerge in the managerial logic of increased efficiency and effectiveness, and my findings reveal that, when we contextualise its practice, a more complex and nuanced understanding emerges. I will use Miller and Rose's (1990) characterisation of governmentality to argue that situating data-driven policing within the police rationale on crime allows me to account for the range of meanings they inscribe onto the constructs of risk and biometric recognition – what I refer to as the organisational optimisation logic of data-driven policing. Miller and Rose (1990) built on the Foucauldian notion of governmentality[20] to distinguish between the 'political rationalities', the discourse that justifies police priorities and actions, and 'technologies of government', the means through which this ideology is translated into action and the interaction between the two (Henman, 2011, 2010; Miller and Rose, 1990). I will first explore the police political rationale to crime, which will serve as my starting point to argue that the organisational optimisation logic manifests itself on both a strategic and on an operational level. Strategically, it allows the police to adhere to a normative construction of what a competent, reliable, and professional police should look like, and operationally, it allows police to justify their desire to expand their nature and scope.

My findings chapters foregrounded that, while there is an understanding that crime results from an unequal distribution of power and structural inequalities in societies, from the perspective of the police, crime is primarily positioned as the result of the flaws of the individual who commits it and of the police who do not prevent it. Dominant scholarly debates often use the term *prevention through prediction* within a spatial context of being at the right place at the right time to deter a criminal event from unfolding. I situate it as an intrinsic part of the police mandate in European society to maintain public order, ensure safety and security, and prevent and investigate crimes, where the logic of prevention is embedded across the organisation. The police political rationale, as such, offers two distinct entry points to crime prevention: modifying or controlling the behaviour of the individuals who commit the crime and investing in the police ability to prevent it. Therefore, I

[20]Governmentality was never published by Foucault but rather developed in his lectures of 1978 and 1979 (Lemke, 2002, p. 50)

argue that solely attributing the construct of risk and recognition to the individual offers a flat ontological view on crime, as it omits the belief that crime, in part, is the result of the police's inability to prevent it. Engaging with both these dimensions allows me to foreground that the constructs of risk and recognition are primarily seen by the police to 'fix' something in the organisation that is not working as well as it should, to justify a certain policing response or crime category, and to affirm the idea of the police as a competent authority. All of these point to a dynamic in which the individual who commits a crime might be the foci of debates on data-driven policing. From the perspective of the police, these functions are primarily directed at overcoming certain organisational challenges. This political rationale, I argue, speaks to an organisational optimisation logic of data-driven policing, which manifests on a strategic and on an operational level.

Strategically, I argue, the use of data-driven policing is believed to reinforce internal and external perceptions of the police as a professional, competent, and reliable authority. In the context of biometric recognition, I found that police have a complicated relationship with the public when it comes to the use of new technologies. Practitioners feel that parts of society will judge the police for not using all possible tools available to prevent or solve a crime. At the same time, it is felt that the use of biometric recognition technologies is subjected to disproportionate levels of public scrutiny. The critiques are seen to overlook police hesitance and care in deploying new technologies and their competence in selecting a reliable computational vendor. This observation is important, as it shows how police perceive public critique on the use of data-driven policing as a direct challenge to the professionalism of the institution and its practitioners. Similarly, in the context of risk scoring, external dynamics, such as the political and media responses to high-profile violence incidents, are seen to quickly shift the blame from the immoral perpetrator to the public institution who 'failed' to prevent the incident from unfolding. In this context, a competent police force is depicted as an authority that prevents violent acts from materialising, placing normative expectations on what the police should look like in contemporary society. Furthermore, the turn to data-driven policing is connected to the police ability to signal to front-line staff that they are part of a professional organisation that is constantly evolving and keeping up with the times. These insights, I argue, show that, strategically, the turn to data-driven policing is closely connected to the productive nature of power; its use is believed to have a significant symbolic value for the authority position of the police, because it is seen to positively influence the internal and external perceptions of the professionalism of the police.

Especially in times of crisis, this productive nature of data-driven policing becomes evident. Most of the use cases I discussed in chapters 5 and 6 emerged from a specific security incident, where the police and external stakeholders believe that the visible use of risk scores and biometric recognition would result in a positive public and political sentiment towards them, again pointing to the interplay between the political rationale and the technologies of governance. The manifestation of a violent incident in part reflects on the inability of the police to prevent it. Visible use of technology signals to the public that the police will do anything to catch the perpetrator and ensure that such a violent act will not happen again in the future – further reinforcing the notion that crime can in part be attributed to the limited nature and scope of the police, which hinders their ability to prevent crime. These insights show that, from the perspective of the police, data-driven policing embodies both the productive and relational dimension of power, and through its visible use, they can negotiate the levels of public trust and confidence in them. This indicates that the turn to data-driven policing is not merely about the belief that it will allow the police to do more with less; rather, it allows them to invoke an image that speaks to the normative construction of what a competent, reliable, and professional police should look like.

Operationally, the police political rationale that views crime in part as a failure of the police to prevent it connects data-driven policing to its perceived ability to fix something in the organisation that is not working as well as it should. Throughout the interviews, practitioners refer to the efficiency affordance of data-driven policing. For example, biometric recognition is seen as a way to automate labour-intensive processes, and data-driven risk scoring should enable police to more accurately select individuals for a specific intervention. This in part validates the dominant theoretical framing of the managerial logic of efficiency and effectiveness, where its use is presented as a way to become more efficient in ‘the task they have set out for themselves’ (Patrick Williams in Richard Vis, 2020). However, my findings illustrate that the perceived organisational benefits of data-driven policing go beyond this intended purpose. In line with studies on the use of predictive policing in Europe (Egbert, 2019; Egbert and Leese, 2021), I found that it justifies the expanding nature and scope of data collection and processing for the purpose of policing. In addition, it creates the conditions for certain new practices to emerge. I will start by outlining the first.

In chapter 5, I explored the different notions of risk, where it enables and justifies a statewide care and control approach. The police and the broader state directly relate the limited success of crime prevention interventions to the inadequate information position and coordination between the police and other public institutions. In this context, data-driven risk scoring is believed to allow the police

to identify the ‘right’ individual and optimise coordination between institutions. Thus, its affordance allows police to optimise for those organisational aspects that are seen to create the conditions for crime to emerge. In the context of recognition, the increased volumes of data available about people and events in a range of public data infrastructure and the existence of data-driven policing create its own policing need. In turn, this need positions more invasive forms of data collection, for example, through social media platforms, CCTV cameras, and telecommunication infrastructures, and different types of analysis, such as biometric comparison to soft biometric classification, as inevitable. This dynamic again points to a belief that crime is the result of the police’s inadequate information position, which justifies their need to expand the nature and scope of evidence collection and further normalises and entrenches the dominant political rationale.

Another operational implication of the police’s political rationale of crime is that emergence of data-driven policing is perceived to open and close avenues for action. This, I argue, has given rise to a number of practices, specifically, intervention, innovation, pilots, transparency, and safeguarding. Here, I will not elaborate on these specific practices, as they have been extensively covered in my findings chapters, but rather use their emergence to argue that the turn to data-driven policing is subjected to a number of internal and external dynamics, which require the organisation to be optimised for its use. Police feel that a number of prerequisites need to be in place for data-driven policing to materialise and become embedded within a large operational organisation like the police force. Take, for example, the limited interest of senior management in technology or the lacking data infrastructures that inhibit contemporary and future data processing. This belief has given rise to a number of practices that are aimed at creating an enabling environment for the organisation for these functions to emerge. In addition, police feel that the emergence of data-driven policing creates a number of negative externalities that need to be mitigated. Take, for example, the insight where police assume that public and political controversy surrounding biometric recognition technologies relates to a lack of information and understanding on the side of the public rather than its actual use or the expanding scope of police power. Or the notion that the ‘just’ use of technologies can be ensured by the creation of procedural and technical safeguards. These observations, I argue, speak to another meaning of the organisational optimisation logic, in which the emergence of data-driven policing not only optimises for specific organisational challenges but also invokes normative expectation about what competent public institutions should look like in the twenty-first century, in which the organisation itself needs to be optimised for the use of data-driven policing functions.

8.1.2. The politics of injustice

A second key finding relates to the politics of injustice, a term I used to account for how the turn to datafication, and more specifically data-driven policing, is invoking new struggles over justice. My research into civic actors' responses to data-driven policing is informed by the recent scholarly debates on data justice (Dencik et al., 2019; Peña Gangadharan and Niklas, 2019; Taylor, 2017). Even though there are different meanings ascribed to this concept, all point to an understanding that the increased datafication of society is a social justice issue that requires new frameworks to understand its impact. Rather than engaging top-down analysis of what 'just' data-driven policing looks like, I draw on Young (2011) and started by listening to new social movements, which I refer to as civic actors, who work on issues of digital rights, human rights, anti-discrimination, and racial justice to gain insights into injustices that emerge from the introduction of data-driven policing. In this section, I will start by outlining what is at stake now that data-driven policing has become part of the logic of contemporary policing – not just in terms of the harms experienced by those affected by this new practice, but also how it entrenches notions of whose voices count in the discussion around police power. To conclude, the concept of the politics of injustice foregrounds that power enacted through data systems raises important questions about the relationship between data-driven policing and how police and police actions are legitimised and legitimacy deficits are formulated.

My findings chapters foreground that civic actors raise a number of injustices in relation to data-driven policing; these primarily pertain to issues of discrimination, criminal justice trap, data protection, governance, and chilling effects. Here, I observe that, while all civic actors engage with questions of police power, there is not one single understanding of what is at stake that would allow them to outline a set of universal principles. Rather, I argue, there are multifaceted meanings and values inscribed into each of these injustices that unmistakably point to three distinct conceptualisations of the loci of struggle: technology as a form of algorithmic governance, the state as the primary granter or rights, and the police as an oppressor. I have also referred to these as the privacy, human rights, and race lenses. Those actors who enter into the debate from an anti-discrimination and racial justice background draw on the lived experiences of targeted communities and critical race theory to position the use of data-driven policing as yet another technology of governance in a long history of oppression and control of Black and Brown communities. Where the police, through its productive power, constructs racialised communities as agents of crime that need to be managed and controlled. In contrast, those actors who approach data-driven policing from a privacy and human rights perspective situate their concerns within a dominant moral philosophy of rights. This approach is closely connected to how justice discourse has been conceptualised within the Westphalian view of nation states (Fraser, 2008) and has been

communicated as universal and critical norms constraining the allocation and exercise of state power (Follesdal, 2005). Here, the state is seen as the primary power holder that grants rights to subjects within their territories, where harms are negotiated with the state within the structures set by them (Hogan, 2011). The nuances lie in that the privacy lens conceptualises injustice primarily in relation to technology and surveillance, and human rights lens in relation to the state not complying with the rights afforded to people in the state–citizen nexus.

In bringing together these, at times conflicting, privacy, human rights, and race lenses, I show how the emergence of data-driven policing has changed the composition, discourse, and actions repertoire of civic actors who challenge police power. This changing composition of resistance explains the emergence of new invisible hierarchies between civic actors' injustice claims in the discussion on police power, a dynamic that I refer to as the politics of injustice. Take, for example, the changing composition of civic actors involved in challenging police power; it brings digital rights advocates and technologists into the discussion on discrimination and policing, and racial and social justice and human rights advocates into the discussion on technology. While, in theory, all challenge police power to create a fairer society, I argue that, in practice, the debates get skewed towards centring data in the discourse on injustice over that of lived experience, racism, and policing. This invokes invisible hierarchies in the articulations of what 'just' or 'unjust' policing looks like and which injustice claims count – a dynamic that is reaffirmed by external stakeholders that privilege data-centric approaches to justice in their funding mechanisms, court rulings, and political discussions. Any injustice frame, I argue, requires an analysis of data-driven policing as a site for struggle that accounts for both the direct harms experienced by those subjected to it and its impact on whose voices count in the discussion on police power.

The politics of injustice further inform how data-driven policing is affecting the action strategies of civic actors, what I refer to as the practice of reform and resistance. In my findings chapter on civic responses to data-driven policing, I highlight how calls for reform are considered a site of conflict. This practice is seen to privilege actions that are aimed at making things less bad by invoking procedural safeguards that will create the conditions for a 'fairer' data-driven policing practice to emerge over those that more fundamentally challenge police power. Which normative expectations of 'fair' are referred to, fairer use of technology or fairer treatment of specific communities, depends on the loci of struggle that inform these reformist practices. The practice of resistance relates to saying no to either the racialised discourse of crime that has become embedded in the data-driven policing practices or drawing redlines on the use of specific technologies. Again, the loci of struggle inform what the normative expectations of *no* are. I will conclude by arguing that

this speaks to the relational nature of police power that is enacted through data systems, in which the use of data systems invokes competing injustice claims that the police can choose to respond to.

My findings on the operational optimisation logic embedded within the police force turn to data and how it invokes a new politics of injustice, which provide novel insights into the actual nature of data-driven policing and the relationship between datafication and police power. These two concepts foreground how the implications of data-driven policing go beyond a mode of governance that tries to manage and control predefined criminal behaviour and requires an analysis of the multitude of ways in which police power becomes embodied, enacted, and challenged through data systems. In the next section, I will theorise about the broader implications of my findings for how we come to understand the struggle over legitimacy and justice in a datafied society.

8.2 Data legitimacy and the negotiation of police power

The role and power of the police in society has never been static. Like the broader state, it has been enabled and subjected to larger changes, such as the Industrial Revolution, the emergence of capitalism, the securitisation of politics, and globalisation. Datafication, scholars argue, is the next paradigm that will significantly shape contemporary politics and economics (Boltanski and Chiapello, 2007; Cohen, 2019; Dencik, 2019; Harvey, 2017; Zuboff, 2015). Each of these paradigm shifts has given rise to a number of man-made crises that have challenged the discourse and norms on governance and justice and who is seen as the legitimate authority (Beck, 1992; Hobolt, 2018). The emergence of a new paradigm has historically led to crises in legitimacy that have shifted the boundaries of how we come to understand power and who we consider to be a legitimate power holder. This understanding of legitimacy as a continuous negotiation between the power holders and the public offers my analysis a framework to theorise about the broader implications of data-driven policing for how we come to understand crime, police power, and justice in contemporary societies. In this next section, I will explore the relationship between data and the legitimacy claim of the police. Here, I build on the premise that the police are not a neutral actor, but rather an agent whose actions are the result of conscious choices that are aimed at justifying and solidifying their power. Civic actors, as alluded to in the previous section, are not a homogeneous audience; rather, their politics of injustice invokes invisible hierarchies on which injustices count. This, in turn, I argue, raises important questions about how we come to understand just and unjust policing in the context of datafied policing.

In chapter 2, I built on debates in the broader social sciences, specifically Weber (1968), who argue that, in democratic societies, power is not exercised through blunt force or the divine right to rule, but needs to be justified. Power is theorised as relational, coercive, normative, and productive in nature. To contribute to knowledge about the relationship between data and police power, I will build on three social sciences frameworks. First, Beetham's (1991b, 1991a) observation that it is not so much about the belief in the police, but to the extent that the police and their actions can be legitimised along the dominant believes and values of a given society. Bottoms and Tankebe's (2012) 'dialogic model of police legitimacy', and its refinements by Martin and Bradford (2021), that identify three spheres in which police power is negotiated: power-holder legitimacy, audience legitimacy, and authorisation of external stakeholders (Bottoms and Tankebe, 2012, 2017; Martin and Bradford, 2021). Third, data justice, not as an ideal notion of justice, but rather as a relational lens through which we can understand power asymmetries in which injustice claims are privileged in our understanding of police power. To conclude, datafication has become an intrinsic part of how legitimacy is negotiated. Therefore, I will put forward the notion of data legitimacy as a relational vector to fully account for the multitude of ways in which data-driven policing is mediating the legitimacy claim of the police.

8.2.1 Datafication and shifting social norms and values of society

Beetham (1991a, 1991b) critiqued Weber for his understanding that legitimacy equals the belief in legitimacy. Rather, he argues that we have to approach legitimacy as a process wherein any claim to power can only be gained and maintained by the extent to which power holders, the public, and other stakeholders can justify it in relation to the dominant social norms and values in society. This observation, I argue, requires an articulation about how social norms come about, for which I draw on Gramsci's notion of cultural hegemony as read by Fink (1988) and Fontana (2008). Gramsci's much-cited work on cultural hegemony preoccupies itself with theorising about the relationship between culture and power within capitalism, in which the interests and power of the political and economic elite is not exercised through blunt force but through cultural controls. Influencing the dominant norms and values of society should allow those in power to impose direction on social life. In turn, decisions that align with these dominant social norms get legitimised through active consent of, not all, but the most important publics (Artz and Murphy, 2000; Fink, 1988). This observation points to the understanding that not everyone has to consent to the exercise of power; it is enough if only important audiences consent. While the notion of cultural hegemony has often been used to understand the interplay between domination and resistance (Adamson, 1983; Fontana,

2008), I will build on Fink's (1988) reading of Gramsci, which argues that hegemony itself should be considered a locus of struggle.

Hegemony points this thesis to the understanding that social norms and values are not static rules; rather, they change through political struggle and paradigm shifts, such as colonialism, capitalism, and globalisation, and these changes have consequences for how society is organised. Viewing the datafication of society as a paradigm shift, as this thesis does, allows me to draw on the debates within the field of media and surveillance studies to understand how it is redefining the norms and values that govern society. As discussed in chapter 2, the growing interest in data-driven policing has to be situated in a long-standing privileging of logical reasoning as a way to manage society (Golumbia, 2009) that inscribes authority to data as the main mode of knowledge production and to those who are able to wield its potential (Van Dijck, 2014). In addition, surveillance theories have pointed to a shift of the disciplinary gaze to the logic of control, in which predictive analytics has become central to the state and the market desire to control the present and future behaviours of people. Scholars in these fields theorise about how the existence of increased volumes of data and technological capabilities in the everyday are shifting social norms and, by extension, create normative expectations on what a reliable and competent organisation and just mode of governance looks like. These insights require any theory of legitimacy to be situated not only in relation to power holder and audience legitimacy but also in relation to the belief in data as a social norms that inscribe meaning into who and what is seen as an authority and how actions are justified along these belief systems.

Before moving on to exploring the three spheres in which police power is negotiated – power-holder legitimacy, audience legitimacy, and authorisation of external stakeholders – I want to draw attention to a dynamic in which this ideology is continuously reinforced by the dominant discourse around data, not only by the scholarly and public debates that laud the possibilities of algorithmic governance but also by those who critique it. Centring data in the struggle over power and justice (Amnesty International, 2020; Van Schendel, 2019) runs the risk of overlooking that it emerges alongside existing ideologies, such as the political rationale of the repressive welfare state (Vonk, 2014) and racialised surveillance (Browne, 2015; Fanon, 1968). This dynamic has been clearly articulated by Duarte: 'critics of recent technological developments may also be blamed for exaggerating the impact of such systems, overlooking the variegated contexts in which they operate and "mythologizing the algorithm's place as an all-powerful arbiter of social life" (Knox, 2016)' (Duarte, 2021, p. 201). This insight requires us to acknowledge that, while datafication is influencing how society is organised alongside historically determined constructs of class, race, and

gender, its impact is magnified to such an extent in scholarly and public debates that it becomes the vector through which we come to understand social, political, and economic changes. This, I argue, further invokes normative expectations on the prominence of data as a medium through which the police engage with the world. Any articulation of the relationship between datafication and police power, as such, has to be cognisant of its unintended impact of centring data as an object of study. I will continue to argue how cultural hegemony in the age of datafication creates an enabling environment that shapes how the police, publics, and other stakeholders can justify or critique the use of data-driven policing practices.

8.2.2 Datafication and power-holder legitimacy

Tankebe (2014a, 2014b), as discussed in chapter 2, argues that the justification of power lies in part within the belief of the power holder that they have the moral right to govern. This belief manifests itself in the discourse, rituals, and norms and values presented by the police, through which they claim to be the legitimate authority and cultivate a culture that justifies their actions (Barker and Mulligan, 2003). Here, I will discuss the relationship between data and power-holder legitimacy, more specifically, the ways in which the visible use of data-driven policing allows the organisation to present certain policing visions to legitimise their existence, what is also referred to as the corporate voice or meso-level legitimacy (Martin and Bradford, 2021). The choice to focus on meso-level legitimacy has been informed by the observation that the majority of data-driven policing tools are still being developed and have not yet been deployed, but are still very much presented as part of an overall policing vision. In this section, I will discuss three distinct manifestations of the power-holder legitimacy in relation to data-driven policing: the visible use of data to maintain their authority in society, the justification of its use on the ‘criminology of the other’ and police actions on the values inscribed onto data, and entrenching procedural justice in what constitutes fair policing. I will conclude this section by arguing that these observations speak to the productive nature of power, in which it is seen to enable police to present a certain image of themselves that reinforces its existence as a legitimate authority in society and justifies certain police priorities, practices, and actions.

When I asked police practitioners why they are interested in specific data-driven policing functions, their response was often one of disbelief at the question and articulated along the lines of, why not them, or if anyone should be able to use it, it is them. This points to a fundamental belief that the police, as the enforcement arm of the state, are the legitimate authority to use data-driven policing to ensure public safety and security in society – what Mulcahy (2013) refers to as the legitimacy

process through representation, in which police normalise specific actions through communicating a particular image. Another common response to my question was that police need to use data-driven policing if they want to keep up or even catch up with the times. This points to the dynamic where the ideological grounds of datafication invokes certain normative expectations on what a professional and competent police organisation should look like in the twenty-first century. A failure to do so is feared to result in declining levels of trust and confidence in their authority in society. The productive dimension of data and police power becomes even more explicit in times of crisis, where I found that the visible use of technology is seen to create an image of a competent police force that will do anything to catch the violent perpetrator and ensure that such a horrific act will not happen again in the future. Thus, from the perspective of practitioners, the datafication of society demands that the police integrate data within their operations to maintain their claim to the moral right to govern.

My findings further highlight that the use of data-driven policing functions themselves are justified by a normative discourse of crime. Take, for example, the case of Catch: its use is legitimised on the basis that it will allow police to catch the 'immoral' perpetrator who looks over the shoulder of a vulnerable senior citizen to uncover their pin code, rob them of their bank card, and withdraw their money. The use of other biometric recognition systems are justified in relation to terrorism, in which events that instil fear in the public imagination are used to legitimise specific practices. The use of risk-scoring models are positioned as a necessary tool to identify those prolific perpetrators who are immune to traditional deterrence and control interventions. This discourse connects the justification of data-driven policing to the literature discussed in chapter 2 on the notion of the criminology of the other (Garland, 2004, 2001). Stereotypes of the 'gang member', 'terrorist', and 'gipsy', position certain criminal acts as intrinsic evil or wicked and nothing like 'normal' upstanding citizens, who cannot be rehabilitated but merely managed and controlled. This discourse justifies increased data collection and police interference on some individuals and communities, where the criminal act that can be controlled through the use of data-driven policing functions are often heavily racialised (Williams, 2015; Williams and Clarke, 2018). Thus, the emergence of data-driven policing is part of a long history of the sense of self and place-making of the institution, in which the use of data is justified by a normative and racialised understanding of crime. In addition, I found that the use of data-driven policing functions justifies the desire to expand the nature and scope of police and their priorities, practices, and actions. McQuillan's (2015) conceptualisation of the 'algorithmic state of exception', where actions being taken on the basis of the output of data create a system of oppression that is outside and yet belonging to the state and its institutions, gives insight into the legitimising effect of data-driven policing. Its use creates the perception that

decisions are informed by neutral, verifiable facts and legitimises an approach and actions taken upon the output, a dynamic that resonated in both my case studies. These ideological grounds of data obfuscate that the who and the what of police interventions, the crime categories prioritised in data-driven policing, and the individuals who are made visible through them are political decisions that are based on a normative understanding of crime. These observations foreground the dialogic relationship between data and the normative discourse on crime, in which the one justifies the other and vice versa.

While these observations on power-holder legitimacy show that the visible use of data justifies police authority and actions, this does not fully account for the police practices that emerge in the wake of the turn to data-driven policing. I will conclude this section by foregrounding that the police responses to the externalities of data-driven policing in turn further solidify how they have historically negotiated the police–public relationship. In the dialogic model of legitimacy, Bottoms and Tankebe (2012) and Martin and Bradford (2021) argue that the justification of power is a relational concept, where the power holder to some extent has to respond to legitimacy deficits raised by different audiences. Procedural justice, top-down constructions of what a justice claim looks like, I argue, have historically allowed police to conform to the normative expectation of what 'fair' and 'just' policing should look like. This dynamic reproduces itself in the rise of the practice of safeguarding and transparency, in which the police believe that the organisation itself needs to be optimised to mitigate the negative externalities that could emerge from data-driven policing – further entrenching the notion of procedural justice at the centre of debates on fair and just policing and closing avenues for other legitimacy deficits.

In this section, I have argued that the datafication of society is affecting the legitimacy claims of the police as a power holder. The distinct ways in which police power becomes operational through the visible use of data, as discussed in this section, speaks to the productive nature of power, in which its use is believed by practitioners to reaffirm its relevance as an authority in society, justify specific actions, and pre-emptively erect practices that allow them to respond to those legitimacy deficits that the organisation can account for.

8.2.3 Datafication and audience legitimacy

In their dialogic model of legitimacy, Bottoms and Tankebe (2012) argue that the police legitimacy claim should be understood as relational, a proposition made by the power holder, that is responded to by different publics, which in some cases requires a response from the power holder. Data-driven

policing as an intrinsic part of the productive nature of the justification of police power and actions, as such, requires us to ask which publics police speak to, who gets to determine what a legitimacy deficit looks like, and what this means for our understanding of justice. This is particularly pertinent as scholars are increasingly connecting the practice of algorithmic governance to generating new conditions of what a person is able to do or be within a certain context (Eubanks, 2017; Jansen et al., 2021; Metcalfe and Dencik, 2019). Again, I start by drawing attention to the fact that the nascent practice of data-driven policing primarily speaks to the meso-level legitimacy claim of the police as an institution and not the micro-level legitimacy claim of the individual police officer on the street. Therefore, I argue that the relationship between data, police power, and legitimacy from an audience perspective needs to be understood in connection to the image that police communicate to the public. In this section, I will discuss the broader implications of the politics of injustice in connection to the dialogic model of legitimacy, specifically, how the multitude of injustices create competing legitimacy deficits that further displace fundamental challenges to police power and the racialisation of crime and what this means for our understanding of and approaches to justice in the context of policing.

As previously mentioned, the justification of data-driven policing aligns with Garland's (2004, 2001) concept of the criminology of the other, in which those people who will be subjected to the algorithmic gaze are framed as intrinsically evil or wicked and nothing like 'normal' upstanding citizens. This is what Amoore (2011) has conceptualised as the 'unknown future threat', 'the contemporary security derivative is not centred on who we are, nor even on what our data say about us, but on what can be imagined and inferred about who we might be' (Amoore, 2011, p. 24). This dynamic speaks to a politics of fear (Garland, 2004), in which messaging about specific violent crimes makes people believe that, at some point, they, too, will become a victim. While these concepts shed light on the general productive nature of power, which assumes a correlation between instilling fear in the public imaginary and audience legitimacy, this offers less insight into the implications of the politics of injustice for our understanding of legitimacy and justice. I turn to the field of criminology and theories on audience legitimacy therein to understand how, through the visible use of data-driven policing, the police speak to and respond to what they consider important audiences.

Empirical criminological research suggests that people's perception of the legitimacy of police power is informed by a wide range of attitudes, beliefs, and experiences, from the public's social and economic contexts to direct and community experiences with the police (Jackson et al., 2012; Antrobus et al., 2015; Owusu-Bempah, 2017; Martin and Bradford, 2021). Meso-level legitimacy

claims are therefore perceived differently by the different publics. Thus, while for some publics, the threat of terrorism might justify more invasive biometric recognition technologies, other publics will see it as a frame that 'legitimizes intrusive racist policing and surveillance' (Williams and Clarke, 2018, p. 1). To understand how different audience perceptions of police actions relate to whose voices are heard and responded to, we need to situate them within the observation that police power is dependent on the active consent of important publics, but not necessarily all publics (Artz and Murphy, 2000; Fink, 1988). My findings chapter on civic responses to data-driven policing shows how different lenses – on privacy, human rights, and race – foreground different injustices, and these intersect with distinct action strategies of reform or resistance. Challenges to police power range from abolitionist frames that place police oppression at the centre of the discussion on injustice to privacy or human rights frames that foreground notions of individual data harms. I will layer the politics of injustice with the dialogic nature of legitimacy to argue that these different sites of struggle create competing legitimacy deficits police can choose to respond to.

Police are strategic agents that select how they respond to the internal and external pressures placed on them. My findings show that their response to competing legitimacy deficits happens in two distinct ways. The police can, to a certain extent, choose to engage with the legitimacy deficits that either centre privacy and human rights over those that are more reformist or abolitionist in nature. I use the term *to a certain extent* to account for other factors that play a role in police responses, primarily endorsement or disapproval of their actions by external stakeholders, which I will discuss in the final part of this chapter. The police responses to public critique on the use of biometric recognition technology show that they primarily respond to the legitimacy deficits that centre technology over those who ask more fundamental questions about their authority and operations. Especially those injustice claims that raise reformist demands on what 'fair' and 'just' policing should look like tend to be repeated and responded to in the emerging police practices of safeguarding and transparency. Thus, I argue that presenting different injustices and demands of what constitutes fair policing creates competing legitimacy deficits that allow police, to a certain extent, to choose what they will respond to. They are inclined to privilege those injustices that do not directly challenge their power and can be mitigated by a process that they know, procedural justice. Therefore, it is important to understand that datafication in relation to audience legitimacy will further displace the question of police power to one of technology and safeguards rather than the normative racialised constructs that are embedded within the understanding of certain crime categories.

The observation of competing legitimacy claims that displace certain voices has implications for how we come to think about justice in contemporary society. As discussed in chapter 2, the concept of legitimacy has historically been tied to procedural justice concerns, which engages with people's assessment of the fairness of police actions according to a set of predetermined rules. In chapter 2, I argued that this approach, which finds its roots in psychology, has primarily engaged with the question of whether people find a specific policing action fair, rather than a more normative understanding of justice, what is fair or what fair procedures should look like, or, more fundamentally, what justice is or should look like in the context of policing and who gets to determine what it looks like. To connect legitimacy more firmly to the normative understandings of justice, I discussed a range of social justice theories in chapter 2. These debates draw our attention to the multitude of meanings that are inscribed into the concept of justice: top-down vs bottom-up, the ideal notion of justice vs a political understanding of justice, and distributive justice vs representation, recognition, restorative, transformational, or data justice. In this research, I drew on Young's (2011) understanding that justice is political, which requires us to unpack how the entry point into the debate on police power inscribes which voices are heard, how we assign value to things, which value and belief systems are privileged in our understanding of society, and who gets to shape the normative notion of the future.

The approach that justice is political shows how the question of which voices are seen to be the authority in formulating a legitimacy deficit is normative and relational, dependent on the positionality of those formulating it and how others interact with it. The recognition and misrecognition of legitimacy deficits, I have argued, are part and parcel of the dialogic nature of legitimacy. This finding informs the broader data justice discussion (Dencik et al., 2019; Taylor, 2017), in which scholars try to account for the ways in which datafication intersects with broader social justice. My findings reveal that the datafication of society is creating normative expectations on how policing ought to be organised and that centring any injustice claim will inadvertently displace another in the discussion on police power. This supports the approach that, rather than formulating one theory of data justice in relation to datafication, we should engage it as a relational lens that situates data in larger questions of power, oppression, and harms (Dencik et al., 2019, 2018). Furthermore, my findings speak to Peña Gangadharan and Niklas's (2019) observation that decentring data and technology in an understanding of justice will offer a more nuanced account of what is at stake. Therefore, I will conclude that data justice as a relational lens allows researchers to reveal and explain both harms experienced by those subjected to data systems and the politics of injustice that result from shifting power relations in society. Data justice, as such, allows us to account for the broader social implications that emerge from the introduction of data-driven

policing, not just the ones that get the most visibility or the ones police are most inclined to respond to. This offers the dialogic model on legitimacy a sociological theory of justice to understand the impact of data systems on perceptions fair and just policing.

8.2.4 Authorisation of police power by other stakeholders

The third and final sphere in the dialogic model is that of external stakeholders and how the police claim to power is dependent on their endorsement or disapproval. My findings chapters reveal how political leaders on an (inter)national and municipality level create financial, political, and even spatial conditions that enable and at times actively endorse data-driven policing. For example, financially, in times of austerity, the Home Office in the UK actively invests in police data infrastructures and their uptake of new technological capabilities in the belief that it will allow the police to do more with less. Spatially, the municipality of Amsterdam's desire to test technologies in so-called living labs created space for police to experiment with the development of a privacy-by-design face comparison system. Politically, responses to high-profile violent events actively endorse and promote different data-driven policing technologies, as it is believed that this will meet a perceived policing need and increase positive public sentiment towards the police and the state. These developments show how external stakeholders, such as ministers, politicians, and oversight bodies are also not immune to the shifting social norms that result from the datafication of society.

Fourcade and Gordon (2020) position the modern nation state as a dataist state by drawing on Harari's (2017) concept of dataism, to show that data affordances are increasingly driving the political rationale of states. 'Dataism is not an inevitable consequence of using data in governance. It is an ideology that finds the purpose of government in what can be measured rather than in the will of the people' (Fourcade and Gordon, 2020, p. 81). Again pointing to the notion the datafication of society is shaping how ministers, politicians, and oversight bodies come to understand the world around them, where data processing is perceived as a superior technology of governance to manage social challenges. To understand how the dataist state, and their focus on managing the effect rather than the causes of social challenges, has become intertwined with the political rationale that governs the relationship between crime, police, and society, I will turn to Beck's (1992) risk society as discussed in chapter 2. Governmental risk approaches are what Beck calls 'procedures in order to secure or repair credibility, without fundamentally questioning the form of power or social control involved' (Beck, 1992). Putting Beck's concept of the risk in conversation with the dataist state sheds light on the dynamic where practices of data-driven policing align with and enable the state's need to manage the consequence of risk without having to challenge historically inscribed power

structures and social inequalities. Data entrench a new mode of governance; it further depoliticises state decision-making by collapsing the political into the technical (Andrejevic, 2020). This broader dynamic justifies the political rationale where crime is seen as a flaw of the individual who commits it and the police failure to prevent it, which creates an enabling environment for the police to turn to data-driven policing; in effect, it also depoliticises the normative and often racialised constructions of crime.

I conclude my analysis by arguing that datafication is changing how we come to understand just and fair policing. I build on theories of legitimacy and social justice to theorise about the multiple ways in which police power becomes embodied, enacted, and contested in data systems. I put forward the concept of data legitimacy to account for the distinct and stratified ways in which data-driven policing is mediating the justification of police power. I found that the ideological grounds of datafication invoke normative expectations on the role of data in governance and policing. External stakeholders, who are also not immune to these shifting norms, create an enabling environment that justifies and at times actively promotes the police's turn to data-driven policing. For practitioners, the visible turn to data allows the police to present a certain organisational vision that reinforces their role in society and justifies certain police priorities, practices, and actions. Civic responses to datafication are diverse and create competing legitimacy deficits police can choose to respond to. These dynamics further displace the question of police power to one of technology rather than how data-driven policing reinforces the normative racialised constructs that are embedded within the understanding of certain crime categories. Therefore, I put forward the concept of data legitimacy as a relational vector to account for how datafication is transforming the lines on which police power is justified and negotiated.

8.3 Conclusion

In this thesis, I set out to study the nature of data-driven policing and the relationship between datafication and police power. Much debate on datafication and surveillance has focused on the affordance of data to modify present and future behaviour of people as well as the new challenges that emerge from this pre-emptive gaze. These theories offer global and abstract insights into the relationship between data, power, and justice. Yet to account for the broader implications of datafication in policing, I argued that there is still a lot unknown about its actual practice in Europe, and the multitude of ways its use becomes intertwined with questions of police power and justice is understudied. In my thesis, I have drawn on social scientists, specifically Weber, Beetham, and Bottom and Tankabe, to foreground that police power is productive, coercive, normative, and

relational in nature. Power in democratic societies is rarely enacted through blunt force or the divine right to rule; instead, it has to be justified. Thus, it is dialogic, a proposition made by a power holder, that is responded to by different publics, which, in some cases, requires a response from the power holder. In this context, I approached data-driven policing as both a practice and a site of struggle, which allowed me to theorise about the broader implication of datafication for how we come to understand police power and justice. The use of interviewing practitioners gave me rare insights into the policing visions behind the introduction of data-driven policing, and how its emergence is changing the way problems, needs, and challenges are defined. Interviewing a broad range of civic actors gave me insight into what is at stake with the police turn to data, not only by understanding the impact on those who are at the receiving end of police actions but how it displaces more fundamental challenges to policing, erecting invisible hierarchies on whose voices count in the public debate on police power.

My main contribution to the scholarly and public debate is the observation that datafication is shifting the lines on which police power is justified and negotiated. The authority given to data processing is shifting social norms and creating an environment that enables the police to turn to data-driven policing. More specifically, it is perceived to create normative expectations on what a professional and competent police force looks like in the twenty-first century. Therefore, the turn to data-driven policing is at least in part productive, as it allows the police to conform to this normative expectation and present a specific policing vision. In addition, its use is seen to justify a control approach to complex social problems, where the affordance of data is creating a need and an opportunity to expand the nature and scope of a conservative security gaze within policing and beyond. Datafication is also changing the terms on which civic actors come to challenge police power, which has implications for how we come to talk about unjust and just policing. The injustice claims that centre privacy and, to some extent, human rights gain traction in the public debate and with the police; in turn, these privacy claims displace more fundamental concerns about the normative and often racialised understandings of crime and police oppression. These insights have broader implications for how we come to understand police power and social justice in contemporary societies, in the sense that datafication is changing the grounds on which police power is justified and challenged. The police are confronted with new normative expectations and their actions are further entrenching historic and often racialised understandings of crime.

9. Conclusion: A study of data-driven policing

The exercise and negotiation of power in society is an ongoing debate in social sciences. Different academic disciplines theorise about how power becomes embodied and enacted within their field of study; media and surveillance studies concerns themselves with the relationship between datafication and power, criminologists with police power, and social justice scholars with justice concerns. My thesis built upon these debates and explored how police power and counter-power becomes embodied and enacted in a datafied society. I draw upon the debates within the field of criminology to show that the role and power of the police in society has never been static and is continuously reinforced, challenged, and negotiated. By situating data-driven policing within the institution's role in society, their turn to data becomes tied up in questions of power and legitimacy. This allowed me to theorise about the broader implications of the datafication of society for how we come to talk about unjust and just policing. Here, I found that the turn to data is part of a legitimising frame that is shifting the lines on which police power is justified and negotiated. In this conclusion, I will briefly outline the debates that informed my research focus and lay out my main argument on data legitimacy as a relational vector to account for the multitude of ways in which data is seen to mediate the relationship between the police, their publics, and the external environment. I conclude with suggestions for future studies in the field of data, police, and justice.

A growing field of study concerns itself with the relationship between data and power. Here, research in the field of media studies has offered ample evidence on the ideological grounds of data, where data is seen as an authoritative voice in contemporary societies (Van Dijck, 2014). This consequently frames who is seen as a rightful power holder (Andrejevic, 2014) and how data is shifting what is seen as the raw material for economic production (Sadowski, 2019). Surveillance scholars have contributed to the understanding of how contemporary data infrastructures allow for those in power to monitor and control the present and future behaviours of people for political and economic purposes (Ball and Snider, 2013; Gandy Jr, 2021; Lyon, 2007; Martin et al., 2009). The act of centring data in the understanding of power has allowed scholars to uncover what is changing and how predictive analytics allows for a new mode of governance. Yet, as I have come to argue, when we draw on social scientist theories of power, specifically Weber (1968), who observed that in democratic society, power is not enforced through blunt force or the divine right to rule but has to be justified, the focus on the functional enactment of power through data systems leaves its relational, normative, and productive dimension under-theorised.

Within the datafication of society (Mayer-Schönberger and Cukier, 2013), I situated the state, a traditional power holder in European societies, as a reflective agent (Jessop, 2016) that continues to adapt and negotiate their mandate, building upon historical techniques of maintaining control (Easterling, 2011; Leese, 2020; Scott, 1998). The state has and continues to work to retain its position as the rightful ruling authority in the face of significant shifts in society, such as the Industrial Revolution, the emergence of capitalism, the securitisation of politics, globalisation, and datafication. Their responses to the datafication of society can be found in the state regulatory responses, investment in the local data economy, and the increased number of projects that try to embed data as mode of governance within statecraft. The latter has resulted in ample public controversies on the state's use of data in Europe. Take, for example, the controversies around the use of discriminatory automated welfare distribution systems in the Netherlands (Roosen, 2020), the discriminatory UK A-level exam algorithm used in 2020 (Metcalfe and Jansen, forthcoming), and police use of automated recognition systems across Europe (Amnesty International, 2020; Big Brother Watch, 2018; Kayser-Bril, 2020; Kind, 2019). These developments suggest that data is increasingly becoming a state practice and a site of struggle.

The police, as the most visible agent of the state, play a significant role in maintaining and reinforcing the state authority and vision of how society ought to look through threat of punishment (Bayley and Shearing, 1996; Bourdieu, 1991; Loader, 1997). Throughout my thesis, I have argued that the police has been one of the leading public authorities to experiment with a variety of data practices (Couchman, 2019; Egbert and Leese, 2021; Williams and Kind, 2019). As such, it offers a particularly pertinent field of study. I referred to these data practices as data-driven policing to encompass the multitude of technological functions police are investing in (Jansen, 2018). I observed that public controversies highlight the police's increased interest in data as a medium to engage with societal problems, and scholarly debates have shed light on the ideological grounds of datafication, but I build upon Brayne and Christin's (2021) argument that there is still a lot unknown about the actual practice of data-driven policing. The study of practice (Couldry, 2004) allows me to examine to what end the police are turning to specific data-driven policing functions, how these materialise in their everyday reality and their understanding of crime, and the ways in which it is opening and closing opportunities for future policing action in different contexts.

I showed that, if we only look at police practices, then we only see a partial view of how data influences the legitimisation frames of the police, as it does not reveal how police power is experienced nor how it is challenged (Bottoms and Tankebe, 2012; Martin and Bradford, 2021). Here, for example, the renewed momentum of the Black Lives Matters protests after the murder of

George Floyd and the public controversies on police use of specific technologies show how the exercise of police power is continuously challenged and, to some extent, responded to by the police. Thus, a study of police power requires both an inquiry into the actual practice of data-driven policing and the responses of different publics. To account for this, I included data from civic actors to explore the responses to new forms of datafied policing. This approach was informed by the ongoing developments in the field of social justice. I specifically built on Young (2011), who argues that there is not one ideal theory of justice, but rather to understand what systems of oppression look like, we need to listen to affected communities and to 'new social movements' to gain insights into the injustices that materialise from the emergence of data-driven policing. My choice to listen to 'new social movements', or what I call civic actors, over the voices of affected communities, was informed by the insights of my mappings chapter and case studies, which highlighted that the actual practice of data-driven policing is ephemeral in nature, where most are still in the development phase and might never be deployed on the streets. Rather, at the moment, it is the logic of data as a medium for control that is primarily challenged by civic actors.

As such, this thesis has explored the relationship between data and police power. I conducted multi-sited empirical research into data-driven policing to uncover the practices of police, civic actors, and relevant stakeholders, which enabled me to move beyond the ideological grounds of datafication, the functional enactment of power through data systems, and the overwhelmingly speculative public debates about the use of data-driven policing (Brayne and Christin, 2021). Consequently, I was able to shed light on their actual practices and account for what is changing when data becomes intertwined with policing and how algorithmic forms of governance transform the negotiation of existing systems of power. Through mapping what was happening in the context of Europe and case studies on data-driven risk scoring, biometric recognition, and civic responses to data-driven policing, I answered my main research questions: what is the nature of data-driven policing? What is the relationship between datafication and police power?

My research was rooted in the critical realist tradition, in which knowledge about society is viewed as socially constructed, contextual, and stratified. This approach requires any theory about society to be grounded within a specific context and, as such, has informed my choice for empirical research on data-driven policing in Europe. In this context, the state and its institutions play a key role in people's everyday lives. My research approach builds on Couldry's (2004) approach of media as practice to gain insights into how these systems are used and perceived. Data was collected through fifty-six semi-structured interviews, which allowed me to listen to the experiences, attitudes, and beliefs of experts, police practitioners, and civic actors, and I

substantiated this data with participant observations from police and civic actors meetings and with the study of grey literature. These methods offered insights into the origins and actual uses of data-driven policing functions and how their emergence is shaping how police come to think about and act on crime. These data collection methods also allowed me to explore how the perceived use of data-driven policing has broader implications for how we come to understand the struggle over legitimacy and justice in the age of datafication.

Critical realism offers its own limitations, as insights are generated through a process of double hermeneutics, where the researcher interprets the interpretations of others. Thus, my findings, to some extent, have been influenced by my positionality, experience, and worldview. In addition, my research department had no prior relationships with police departments, which required me to reach out coldly, or through existing networks, to police practitioners with a request for interviews. This will have skewed the practitioner sample to primarily include those voices who were open to discussing police matters with external researchers. Finally, the choice to do multi-sited empirical research, where I looked across and between the implementation of risk scoring and biometric recognition across jurisdictions, allowed me to identify broad organisational principles that structure police approaches to technology and offer insights into the social structures that (re)produce them. However, this approach lacked the breadth and depth to offer a comparative analysis between the distinct deployments. In the third and final section of this chapter, I will offer suggestions for future research that, in part, can overcome these limitations and could reaffirm or dispel my findings.

My findings reveal and explain that the datafication of society is affecting struggles over just and unjust exercise of police power in a number of ways. I first set out to understand what was happening in the context of Europe. My mapping, chapter 4, revealed that, while data-driven policing is a nascent practice and the projects are often ephemeral in nature, police forces are increasingly interested in it. However, the extent to which the police invest in different data-driven policing functions is very much dependent on the organisational cultures, structures and budgets, and the wider political climate within which they operate. Their turn to data is in some contexts enabled, actively promoted, and justified by external stakeholders, such as ministers, politicians, oversight bodies, and other public institutions. These developments informed my choice that, when looking at data as practice, it is imperative to move beyond the isolated technical artefact or a specific project, as it might never materialise in practice, and research the emergence of a specific function within policing.

My case studies, chapter 5 and 6, spoke to my research question, as they allowed me to describe the actual practices of data-driven risk scoring and biometric recognition and the political rationale that underpins the introduction of these practices. I have shown how, from the perspective of the police, there are two distinct entry points into crime prevention: managing the individual who commits it and improving police operations. The latter emerges from the political rationale that crime, in part, is the result of a police failure to prevent it. Here, I found that the turn to data-driven policing primarily embodies the police's need to 'fix' something that is not working as well as it should within policing; as such, I argued that solely attributing the construct of risk and recognition to the 'at-risk' individual offers a flat ontological view. These insights offer nuances to the dominant understanding of to what end the constructs of risk and recognition are deployed. To account for these findings, I introduced the term *organisational optimisation logic* to distinguish between the managerial logic of increased efficiency and effectiveness often attributed to the introduction of algorithmic governance systems and the productive and normative dimension of police power that becomes embodied and enacted through data systems. This is important as most scholarly and public debates focus on the functional enactment of police power through data systems, yet my findings reveal how the emergence of data-driven policing allows police to adhere to the normative expectations that the datafication of society places upon a public authority. In turn, the emergence of data-driven policing further entrenches normative notions of crime and crime prevention in the operations of policing, in which the perceived neutrality of crime insights based on data-driven decision-making justifies certain actions and displaces discussions on the racialisation of crime to questions of under which organisational, ethical, and legal conditions police can use it.

Approaching the emergence of data-driven policing as a site of struggle, chapter 7, revealed that there is not a single understanding of what is at stake with its introduction. Rather, the emergence of data-driven risk scoring and biometric recognition creates a politics of injustice, as it invokes invisible hierarchies between which voices are privileged in the debate on police power. The interest in focusing on the new data systems runs the risk of magnifying the impact of data-driven policing to such an extent that it becomes the vector through which we come to understand just and unjust policing. This has broader implications for how we come to understand police power and justice in contemporary societies. Historically, the justification of police power has been tied to the concept of procedural justice, which engages with people's assessment of the fairness of police actions according to a set of predetermined rules that are defined by those who are the least affected by police oppression. Building on Young's (2011) understanding that justice is political, I argued that listening to civic voices on what is at stake reveals that the datafication of society is shifting the boundaries of how we come to understand just policing. The more reformist and data-centred

actions and arguments are increasingly shaping the public debate and offer low-hanging legitimacy deficits police can respond to. This, in turn, was believed to obfuscate and displace those agents and arguments that try to expose and challenge the historically inscribed racialised notions of crime and the failure of the police to keep communities of colour safe. Therefore, I build on scholars (Dencik et al., 2019, 2018) who position data justice as a relational lens to argue that any inquiry into the impact of datafication on society needs to account for the harms experienced by those affected by this new practice, but also how it entrenches notions of whose voices count in the discussion around police power.

As alluded to in the introduction of my conclusion, the role and power of the police in society has never been static, and in my analysis, chapter 8, I situated their turn to data within their broader legitimacy claim, which is continuously negotiated. I put forward the concept of *data legitimacy* to account for the distinct and stratified ways in which the datafication of society is seen to mediate and justify police power. I found that the ideological grounds of data have placed normative expectations on what contemporary policing should look like, which manifested in the relationship between the police, the public, and external stakeholders. Practitioners believe that public trust and confidence in the police is dependent on their ability to keep up with the times and view the visible use of data functions as a way to communicate a policing vision that reinforces their standing as a competent, reliable, and professional authority. Here, the police turn to data does not operate in isolation, but rather I found that ministers, politicians, and other stakeholders create an enabling environment and at times actively endorse the use of data-driven policing functions. Civic responses to these developments create competing legitimacy deficits that police can choose to respond to. As such, the recognition and misrecognition of legitimacy deficits, I argue, are part and parcel of the dialogic nature of legitimacy. These insights substantiate my main argument that datafication has become an intrinsic part of how police power is justified and legitimacy is negotiated.

9.1 Contributions and future directions

In the remainder of this chapter, I will outline the implications of my research for academic, policy, and practitioner debates and make suggestions for future studies in the field of data, police, and justice. My multi-sited empirical research approach into data as practice offers an in-depth view on the use and perception of data-driven policing, which informs broader academic debates in a number of ways. My findings have theoretical implications for media and surveillance studies, as they contributes to theory building on the relationship between data and power. Here, I moved beyond the ideological grounds of datafication and the ways in which it enacts mediums of control

and social sorting to account for the productive value of data in the justification of historically determined systems of power. This could explain why the police and the broader state are visibly investing in overwhelmingly speculative data systems, despite disappointing results and public critiques. It contributes to criminology and social justice debates, as it foregrounds how datafication creates a normative expectation of what just and fair policing looks like. In addition, this research forms an integral part of the broader data justice project funded by the European Research Council, which aims to understand the implication of the datafication of police, migration, and low-wage labour in relation to social justice. My findings contribute to the project's understanding of data justice, not as an ideal notion of justice but rather a relational lens through which we can understand the relationship between data and social justice.

The act of contextualising has been at the centre of my research approach. I argued that contextualisation is imperative to fully account for the material effect of datafication on the police. Generalising findings from one context to another runs the risk of reducing a more nuanced understanding of the social, political, and organisational context that shape and are shaped by the emergence of algorithmic governance to a common denominator, data systems. This overlooks the specificity of how datafication becomes intertwined within historic and ongoing struggles for how society is organised. Situated research, as such, offers insights into the specific political rationale and organisational interests that are embodied and enacted within data systems. The act of contextualising also makes it unclear to what extent my research findings can be generalised to other contexts inside and outside Europe without losing the nuances that allowed me to account for data legitimacy as a relational vector that mediates the relationship between the police, their publics, and external stakeholders. I would therefore recommend future research into data legitimacy to focus on how it manifests in Eastern or southern European countries, or even in other contexts. Furthermore, I primarily engaged with meso-level legitimacy, as most of the data-driven policing functions were not yet deployed on the street. Another interesting research angle could be how data legitimacy mediates the justification of power on the streets, the micro-level legitimacy claims.

Research on data-driven policing as a practice and a site for struggle is highly relevant for the European policy context. In the European policy arena, debates on data are informed by the European Commission (EC) and Member States' beliefs that economic progress and social well-being is increasingly dependent on the data economy. To ensure that Europe does not fall behind and that it reaps the benefits of the next-generation data economy, the EC is heavily investing in and regulating the data landscape (European Commission, 2020; Jansen, 2021). European public institutions are seen as key actors in shaping Europe's data future. The EC observed that the

consumer data market is dominated by non-European companies outside the control of Europe, yet the large volumes of data about people, objects, and society held by public authorities are believed to offer new opportunities for an industrial algorithmic market (European Commission, 2020). These same institutions are seen by the EC as possible consumers for home-grown European algorithmic governance products. European policy debates, as such, invoke a normative understanding of what contemporary public authorities should look like and create an enabling environment that promotes the state's turn to data (Jansen, 2021). My findings inform these policy debates by offering a more nuanced account of what is at stake with the datafication of the state; rather than encouraging the experimentation on people that are dependent on public authorities for their livelihoods and trying to mitigate some of datafication's externalities through safeguarding, it is imperative to proceed with caution and listen to a those who will be most affected.

Recent public debates have centred the use of data in the controversies around oppressive and harmful state programmes (Amnesty International, 2021; Roosen, 2020), even though, in these cases, data often only played a small part in the structural violence that is enacted through the bureaucracy of the state. My research revealed that, while this framing resonates with public and scholarly interests in what is new, where data is treated as exceptional, it deflects the conversation from what is actually at stake, racialised and oppressive policing practices. The violence that is rendered through the politics of injustice becomes painfully clear in the observations of those who experience and fight police oppression, who said that centring data in the struggle over injustice and police power made them feel like outsiders in their own struggles. These insights have theoretical and civic implications, as that focus on the new runs the risk of attributing harms to data systems that are the direct result of repressive politics and structures of oppression. To more accurately account for the implications of datafication, any study on data harms should engage with data-driven systems as a site of struggle, which enacts violence on those that are subjected to it and those who try to challenge the dominant social norms and values that normalise and justify its use on resource-poor and marginalised communities. Furthermore, civic actors should actively reflect on how their actions offer competing injustice claims that determine whose voices are included in policy discussions that shape what our futures will look like, as it might negatively impact others who are operating in a similar space. The limitations of my multi-sited research were that I chose to engage with the responses of civic actors over the experience of targeted communities. Recommendations for further research on the politics of injustice are to focus on those data-driven models that are already part of everyday policing activities and to listen to the voices of the communities who are subjected to it.

References

Abdelwhab, A. and Viriri, S. (2018) 'A survey on soft biometrics for human identification', in Yang, J. et al. (eds) *Machine Learning and Biometrics*. London: IntechOpen, pp. 37–56.

Abraham, M. *et al.* (2011) *Pilots ProKid Signaleringsinstrument 12- geëvalueerd*. Amsterdam, The Netherlands: DSP-groep, pp. 1–90. Available at: https://repository.wodc.nl/handle/20.500.12832/1832 (Accessed: 15 March 2021).

Access Now *et al.* (2014) *Necessary & Proportionate: International Principles on the Application of Human Right to Communications Surveillance*. Available at: https://necessaryandproportionate.org/images/np-logo-og.png (Accessed: 13 December 2021).

Access Now (2021) *EU takes minimal steps to regulate harmful AI systems, must go further to protect fundamental rights*, 21 April. Available at: https://www.accessnow.org/eu-minimal-steps-to-regulate-harmful-ai-systems/ (Accessed: 17 August 2021).

Ackroyd, S. and Karlsson, J.C. (2014) 'Critical realism, research techniques, and research designs.', in Edwards, P.K., O'Mahoney, J., and Vincent, S. (eds) *Studying organizations using critical realism: A practical guide*. Oxford: OUP, pp. 21–45.

Adamson, W., L. (1983) *Hegemony and revolution: A study of Antonio Gramsci's political and cultural theory*. Berkeley, Los Angeles, and London: University of California Press.

Aktinson, P. and Hammersley, M. (1998) 'Ethnography and participant observation.', in *Strategies of Inquiry.* Thousand Oaks; London; New Delhi: Sage Publications.

Algemene Rekenkamer (2016) *ICT politie 2016*. Den Haag, The Netherlands: Algemene Rekenkamer. Available at: https://www.rekenkamer.nl/publicaties/rapporten/2016/12/13/ict-politie-2016 (Accessed: 1 December 2021).

Amnesty International (2018) *Trapped in the Matrix: Secrecy, stigma, and bias in the Met's Gangs Database*. London, UK: Amnesty International. Available at: www.amnesty.org.uk/gangs (Accessed: 4 August 2019).

Amnesty International (2020) *We sense trouble: Automated discrimination and mass surveillance in predictive policing in the Netherlands.* EUR 35/2971/2020. London, UK: Amnesty International. Available at: https://www.amnesty.org/en/documents/eur35/2971/2020/en/ (Accessed: 15 March 2021).

Amoore, L. (2011) 'Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times', *Theory, Culture & Society*, 28(6), pp. 24–43.

Amoore, L. (2020) *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham: Duke University Press.

Amsterdam (2021) *Digitale Perimeter*. Amsterdam, The Netherlands: Gemeente Amsterdam. Available at: https://www.amsterdam.nl/wonen-leefomgeving/innovatie/de-digitale-stad/digitale-perimeter/ (Accessed: 26 March 2021).

Ananny, M. and Crawford, K. (2018) ‘Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability’, *New Media & Society*, 20(3), pp. 973–989. doi:10.1177/1461444816676645.

Andrejevic, M. (2012) ‘Exploitation in the data mine’, in Fuchs, C. et al. (eds) *Internet and surveillance; The Challenges of Web 2.0 and Social Media*. New York: Routledge, pp. 71–88.

Andrejevic, M. (2013) *Infoglut: How Too Much Information Is Changing the Way We Think and Know.* London and NY: Routledge.

Andrejevic, M. (2014) ‘Big Data, Big Questions| The Big Data Divide’, *International Journal of Communication*, 8(17), pp. 1673–1689.

Andrejevic, M. (2020) *Automated Media*. New York and Abingdon: Routledge.

Ansorge, J.T. (2016) *Identify and sort: how digital power changed world politics*. New York: Oxford University Press.

Antrobus, E. *et al.* (2015) ‘Community norms, procedural justice, and the public’s perceptions of police legitimacy’, *Journal of contemporary criminal justice*, 31(2), pp. 151–170.

Archer, M.S. (1995) *Realist social theory: The morphogenetic approach*. Cambridge, New York and Melbourne: Cambridge University Press.

Archer, M.S. (ed.) (2013) *Social morphogenesis*. Dordrecht, Heidelberg, New York and London: Springer Science & Business Media.

Artz, L. and Murphy, B.O. (2000) *Cultural hegemony in the United States*. Thousand Oaks; London; New Delhi: Sage Publications.

Auger, C.P. (1998) *Information Sources in Grey Literature. Guides to Information Sources*. 4th edn. London: De Gruyter Saur.

Bain, A. (2016) ‘Horses and Horsepower, Fingerprints and Forensics; The Development of Technology and Law Enforcement’, in Bain, A. (ed.) *Law enforcement and technology: Understanding the use of technology for policing.* London, UK: Palgrave Pivot.

Ball, K. (2010) ‘Workplace surveillance: an overview’, *Labor History*, 51(1), pp. 87–106. doi:10.1080/00236561003654776.

Ball, K., Haggerty, K. and Lyon, D. (eds) (2012) *Routledge handbook of surveillance studies*. Abingdon and New York: Routledge.

Ball, K. and Snider, L. (eds) (2013) *The surveillance-industrial complex: A political economy of surveillance*. Abingdon and New York: Routledge.

Baraniuk, C. (2018) ‘Exclusive: UK police wants AI to stop violent crime before it happens’, *New Scientist*. Magazine issue 3206, 26 November. Available at: https://www.newscientist.com/article/2186512-exclusive-uk-police-wants-ai-to-stop-violent-crime-before-it-happens/ (Accessed: 13 July 2021).

Barbrook, R. and Cameron, A. (1996) ‘The californian ideology’, *Science as Culture*, 6(1), pp. 44–72.

Barker, R. and Mulligan, S.P. (2003) ‘Legitimating identities: the self-presentations of rulers and subjects’, *Canadian Journal of Sociology*, 28(4), pp. 572–575.

Bastos, F.B. and Curtin, D.M. (2020) ‘Interoperable information sharing and the five novel frontiers of EU governance: a special issue’, *European Public Law*, 26(1), pp. 59–70.

Bayley, D.H. and Shearing, C.D. (1996) ‘The future of policing’, *Law & Society Review*, 30(2), pp. 585–606.

Beck, U. (1992) *Risk Society: Towards a New Modernity*. Translated by M. Ritter. London, Newbury Park and New Delhi: Sage Publications.

Beckford, M. (2018) ‘Police computer “predicts” who is likely to commit crime’, *Mail Online*, 23 December. Available at: https://www.dailymail.co.uk/news/article-6523997/Minority-Report-police-computer-predict-likely-commit-crimes.html (Accessed: 15 March 2021).

Beetham, D. (1991a) ‘Max Weber and the legitimacy of the modern state’, *Analyse & Kritik*, 13(1), pp. 34–45.

Beetham, D. (1991b) *The legitimation of power*. London: Macmillan Press.

Benjamin, R. (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*. Medford MA: Polity Press.

Berman, G., Coleman, C. and Taylor, M. (2012) *Police and crime commissioner elections, 2012*. London: House of Commons Library. Available at: https://researchbriefings.files.parliament.uk/documents/RP12-73/RP12-73.pdf (Accessed: 2 February 2019).

Bhaskar, R. (2008) *A realist theory of science*. Abingdon and New York: Routledge.

Big Brother Watch (2018) *Face Off: The lawless growth of facial recognition in UK policing*. London: Big Brother Watch. Available at: https://bigbrotherwatch.org.uk/campaigns/stop-facial-recognition/report/ (Accessed: 14 June 2019).

Big Brother Watch (2019) *Regulation would be a life raft for live facial recognition — we need a ban*, *Big Brother Watch*. Available at: https://bigbrotherwatch.org.uk/2019/08/regulation-would-be-a-life-raft-for-live-facial-recognition-we-need-a-ban/ (Accessed: 6 April 2021).

Bigo, D. (2006) ‘Security, exception, ban and surveillance’, in Lyon, D. (ed.) *Theorizing Surveillance: The Panopticon and beyond*. Devon and Oregon: Willan Publishing.

Bigo, D. (2008) ‘Globalized (in) security: the field and the ban-opticon’, in Bigo, D. and Tsoukala, A. (eds) *Terror, Insecurity and Liberty; Illiberal Practices of Liberal Regimes after 9/11*. London: Routledge, pp. 20–58.

Bigo, D. and Guild, S. (2005) ‘Introduction: Policing in the Name of Freedom’, in Bigo, D. and Guild, S. (eds) *Controlling Frontiers: Free Movement Into and Within Europe*. Aldershot and Burlington: Ashgate

Boltanski, L. and Chiapello, È. (2007) *The new spirit of capitalism*. London: Verso.

Bottoms, A. and Tankebe, J. (2012) 'Beyond procedural justice: A dialogic approach to legitimacy in criminal justice', *The Journal of Criminal Law and Criminology*, 102(1), pp. 119–170.

Bottoms, A.E. and Tankebe, J. (2017) 'Police Legitimacy and the Authority of the State', in du Bois-Pedain, A., Ulväng, M., and Asp, P. (eds) *Criminal Law and the Authority of the State.* Oxford, Portland: Hart Publishing Limited.

Bourdieu, P. (1991) *Language and symbolic power*. Edited by J.B. Thompson. Translated by G. Raymond and M. Adamson. Cambridge MA: Harvard University Press.

Boyatzis, R.E. (1998) *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks; London; New Delhi: Sage Publications.

boyd, D. and Crawford, K. (2012) 'CRITICAL QUESTIONS FOR BIG DATA', *Information, Communication & Society*, 15(5), pp. 662–679.

Boyd, D. and Crawford, K. (2011) 'Six Provocations for Big Data', in *A Decade in Internet Time*. *Symposium on the Dynamics of the Internet and Society*, Rochester, NY: Social Science Research Network. doi:10.2139/ssrn.1926431.

Boyne, R. (2000) 'Post-Panopticism', *Economy and Society*, 29(2), pp. 285–307. doi:10.1080/030851400360505.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative research in psychology*, 3(2), pp. 77–101.

Brayne, S. (2017) 'Big Data Surveillance: The Case of Policing', *American Sociological Review*, 82(5), pp. 977–1008.

Brayne, S. and Christin, A. (2021) 'Technologies of crime prediction: The reception of algorithms in policing and criminal courts', *Social Problems*, 68(3), pp. 608–624.

Brayne, S., Rosenblat, A. and Boyd, D. (2015) *Predictive policing*. datacivilrights.org, pp. 3–11. Available at: https://www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf.

Brennen, B.S. (2014) *Qualitative research methods for media studies.* New York: Routledge.

Browne, S. (2015) *Dark matters: On the surveillance of blackness*. Durham: Duke University Press.

Brusseau, J. (2020) 'Deleuze's Postscript on the Societies of Control: Updated for Big Data and Predictive Analytics', *Theoria*, 67(164), pp. 1–25.

Brynjolfsson, E. and McAfee, A. (2014) *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York and London: WW Norton & Company.

Buolamwini, J. and Gebru, T. (2018) 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency in Machine Learning Research*, pp. 77–91.

Cardiff University (2019) *Research Integrity and Governance Code of Practice*. Cardiff, UK: Cardiff University. Available at: https://www.cardiff.ac.uk/__data/assets/pdf_file/0004/937021/Research-Integrity-and-Governance-Code-of-Practice-v3-PDF.pdf (Accessed: 30 November 2021).

Cath, C. and Jansen, F. (2021) 'Dutch Comfort: The limits of AI governance through municipal registers', *arXiv preprint arXiv:2109.02944* [Preprint].

Chowdhury, H. (2018) 'Kent Police stop using crime predicting software', *The Telegraph*, 27 November. Available at: https://www.telegraph.co.uk/technology/2018/11/27/kent-police-stop-using-crime-predicting-software/ (Accessed: 7 October 2021).

Cinnamon, J. (2017) 'Article Social Injustice in Surveillance Capitalism', *Surveillance & Society*, 15(5), pp. 609–625.

Clarke, R. (1988) 'Information technology and dataveillance', *Communications of the ACM*, 31(5), pp. 498–512.

COC (2019) *Tussentijds rapport met corrigerende maatregel betreffende de visitatie bij de Federale Politie van de luchthaven Zaventem door het controleorgaan op de politionele informatie met betrekking tot het gebruik van gezichtsherkenning op de nationale luchthaven van Zaventem*. DIO19005. Brussels, Belgium: Controleorgaan op de politionele informatie. Available at: https://www.controleorgaan.be/files/DIO19005_Onderzoek_LPABRUNAT_Gezichtsherkenning_Publiek_N.PDF (Accessed: 5 October 2021).

Cohen, J.E. (2019) *Between Truth and Power*. New York: Oxford University Press.

Collins, P.H. (1990) 'Black feminist thought in the matrix of domination', *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*, 138(1990), pp. 221–238.

Coole, D. (2005) 'Rethinking Agency: A Phenomenological Approach to Embodiment and Agentic Capacities', *Political Studies*, 53(1), pp. 124–142. doi:10.1111/j.1467-9248.2005.00520.x.

Costanza-Chock, S. (2018) 'Design justice: Towards an intersectional feminist framework for design theory and practice', in *Proceedings of the Design Research Society*. *Catalyst*, University of Limerick.

Costanza-Chock, S. (2020) *Design justice: Community-led practices to build the worlds we need*. Cambridge MA, London: The MIT Press.

Couchman, H. (2019) *Policing by Machine: predictive policing and the threats to our rights.* London, UK: Liberty. Available at: https://www.libertyhumanrights.org.uk/issue/policing-by-machine/ (Accessed: 12 July 2021).

Couldry, N. (2004) 'Theorising media as practice', *Social semiotics*, 14(2), pp. 115–132.

Crawford, K. (2018) *You and AI – Machine learning, bias and implications for inequality*. London, UK: The Royal Society 17 July. Available at: https://royalsociety.org/science-events-and-lectures/2018/07/you-and-ai-equality/ (Accessed: 25 May 2021).

Danermark, B. *et al.* (2002) *Explaining Society: An Introduction to Critical Realism in the Social Sciences.* Abingdon and New York: Routledge.

Danermark, B. *et al.* (2019) *Explaining Society: Critical Realism in the Social Sciences*. 2nd edn. Abingdon and New York: Routledge.

Dantcheva, A., Elia, P. and Ross, A. (2015) 'What else does your biometric data reveal? A survey on soft biometrics', *IEEE Transactions on Information Forensics and Security*, 11(3), pp. 441–467.

De Hert, P. and Gutwirth, S. (2006) 'Interoperability of police databases within the EU: An accountable political choice?', *International Review of Law, Computers & Technology*, 20(1–2), pp. 21–35. doi:10.1080/13600860600818227.

De Hert, P. and Sajfert, J. (2018) 'The role of the data protection authorities in supervising police and criminal justice authorities processing personal data', in Brière, C. and Weyembergh, A. (eds) *The needed balances in EU Criminal Law; Past present and future*. Oxford and Portland, Oregon: Hart Publishing, pp. 243–255.

De Roo, K. (2016) *De politie op weg naar een nieuwe datacenterstrategie*. Available at: https://datacenterworks.nl/artikelen/de-politie-op-weg-naar-een-nieuwe-datacenterstrategie (Accessed: 4 October 2021).

Dean, K., Joseph, J. and Norrie, A. (2005) 'New essays in critical realism', *New Formations*, 56(56), pp. 7–26.

DeKeseredy, W. and Dragiewicz, M. (2018) 'Introduction Critical criminology: Past, present, and future', in DeKeseredy, W.S. and Dragiewicz, M. (eds) *Routledge handbook of critical criminology (2nd Edition)*. Abingdon and New York: Routledge, pp. 1–12.

Deleuze, G. (1992) 'Postscript on the Societies of Control', *October*, 59, pp. 3–7. doi:http://www.jstor.org/stable/778828.

Den Hengst-Bruggeling, M. (2013) *Informatiegestuurd politiewerk*, Politieacademie. Available at: https://www.politieacademie.nl/thema/gebiedsgebondenpolitie/canonggp/planmatigenmethodischwerken/Paginas/Informatiegestuurd-politiewerk.aspx (Accessed: 4 October 2021).

Dencik, L. *et al.* (2019) 'Exploring Data Justice: Conceptions, Applications and Directions', *Information, Communication & Society*, 22(7), pp. 873–881.

Dencik, L. (2019) 'Situating practices in datafication–from above and below', in Stephansen, H.C. and Treré, E. (eds) *Citizen media and practice: Currents, connections, challenges.* Abingdon and New York: Routledge, pp. 243–255.

Dencik, L. (2021) 'The Datafied Welfare State: A Perspective from the UK', in Hepp, A., Jarke, J., and Kramp, L. (eds) *The Ambivalences of Data Power: New perspectives in critical data studies.* Hampshire and New York: Palgrave Macmillan.

Dencik, L., Hintz, A. and Cable, J. (2016) 'Towards data justice? The ambiguity of anti-surveillance resistance in political activism', *Big Data & Society*, 3(2), pp. 1–11. doi:10.1177/2053951716679678.

Dencik, L., Hintz, A. and Carey, Z. (2018a) ‘Prediction, pre-emption and limits to dissent: Social media and big data uses for policing protests in the United Kingdom’, *New Media & Society*, 20(4), pp. 1433–1450. doi:10.1177/1461444817697722.

Dencik, L., Jansen, F. and Metcalfe, P. (2018b) ‘A conceptual framework for approaching social justice in an age of datafication’, *Data Justice Project*. Available at: https://datajusticeproject.net/wp-content/uploads/sites/30/2018/11/wp-conceptual-framework-datajustice.pdf (Accessed: 30 August 2018).

Dencik, L. and Stevens, S. (2021) ‘Regimes of justification in the datafied workplace: the case of hiring’, *New Media & Society* [Preprint]. doi:10.1177/14614448211052893.

Digital Freedom Fund (2021) *Decolonising Digital Rights*. Available at: https://digitalfreedomfund.org/decolonising/ (Accessed: 13 June 2021).

Duarte, D.E. (2021) ‘The The Making of Crime Predictions: Sociotechnical Assemblages and the Controversies of Governing Future Crime’, *Surveillance & Society*, 19(2), pp. 199–215.

Easterling, K. (2011) *Extra-Statecraft: The Power of Infrastructure Space*. London and New York: Verso.

Easton, G. (2010) ‘Critical realism in case study research’, *Industrial Marketing Management*, 39(1), pp. 118–128. doi:10.1016/j.indmarman.2008.06.004.

EDRi (2021) *EDRi submits response to the European Commission AI adoption consultation*. Available at: https://edri.org/our-work/edri-submits-response-to-the-european-commission-ai-adoption-consultation/ (Accessed: 9 November 2021).

Egbert, S. (2019) ‘Predictive Policing and the Platformization of Police Work’, *Surveillance & Society*, 17(1/2), pp. 83–88.

Egbert, S. and Leese, M. (2021) *Criminal Futures: Predictive Policing and Everyday Police Work*. London; New York: Taylor & Francis. Available at: https://library.oapen.org/handle/20.500.12657/42895 (Accessed: 15 March 2021).

Elder-Vass, D. (2010) *The causal power of social structures: Emergence, structure and agency*. New York: Cambridge University Press.

Elmer, G. (2012) ‘Panopticon-discipline-control’, in Ball, K., Haggerty, K., and Lyon, D. (eds) *Routledge handbook of surveillance studies*. Abingdon and New York: Routledge, pp. 21–29.

Engine Room (2016) ‘Responsible Data Handbook’. The Engine Room. Available at: https://the-engine-room.github.io/responsible-data-handbook/ (Accessed: 30 November 2021).

Equinox (2021) *Towards Climate Justice: Rethinking the European Green Deal from a racial justice perspective*. Available at: https://www.equinox-eu.com/wp-content/uploads/2021/06/Towards-Climate-Justice-Equinox.pdf (Accessed: 30 August 2021).

Ethics Committee (2020) *Ethics Committee Minutes and Advice*. West Midlands: West Midlands Police Crime Commissioner. Available at: https://www.westmidlands-pcc.gov.uk/wp-content/uploads/2020/02/17012020-EC-Minutes-and-Advice.pdf?x56534.

Etikan, I. (2016) 'Comparison of Convenience Sampling and Purposive Sampling', *American Journal of Theoretical and Applied Statistics*, 5(1), pp. 1–4. doi:10.11648/j.ajtas.20160501.11.

Eubanks, V. (2017) *Automating Inequality: How high-tech tools profile, police, and punish the poor.* New York: St. Martin's Press.

European Commission (2018) *Final Report Summary - SIIP (Speaker Identification Integrated Project), Cordis*. Available at: https://cordis.europa.eu/project/id/607784/reporting (Accessed: 17 June 2021).

European Commission (2020) *White Paper on Artificial Intelligence - A European Approach to Excellence and Trust*. Brussels, Belgium: European Commission. Available at: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (Accessed: 17 August 2021).

Fanon, F. (1968) *Black Skin, White Masks*. Translated by C.L. Markmann. New York: Grove press.

Fay, B. (1987) *Critical social science: Liberation and its limits*. Oxford: Polity Press.

Ferguson, A.G. (2012) 'Predictive policing and reasonable suspicion', *Emory law journal*, 62(2), p. 259.

Ferguson, A.G. (2017) *The Rise of Big Data Policing, The Rise of Big Data Policing*. New York: New York University Press.

Ferguson, A.G. (2019) 'Predictive Policing Theory', in Lave, T.R. and Miller, E.J. (eds) *The Cambridge Handbook of Policing in the United States*. 1st edn. Cambridge: Cambridge University Press, pp. 491–510. doi:10.1017/9781108354721.025.

Fink, L. (1988) 'The new labor history and the powers of historical pessimism: Consensus, hegemony, and the case of the Knights of Labor', *The Journal of American History*, 75(1), pp. 115–136.

Fiske, J. (1998) 'Surveilling the city: whiteness, the black man and democratic totalitarianism', *Theory, Culture & Society*, 15(2), pp. 67–88.

Flensburg, S. and Lomborg, S. (2021) 'Datafication research: Mapping the field for a future agenda', *New Media & Society* [Preprint]. doi:10.1177/14614448211046616.

Focant, J. *et al.* (2012) *Study on possible ways to enhance efficiency in the exchange of police records between the Member States by setting up a European Police Records Index System*. Brussels, Belgium: DG Home - European Commission. Available at: https://ec.europa.eu/home-affairs/sites/default/files/e-library/documents/policies/police-cooperation/general/docs/epris-final_report_en.pdf (Accessed: 12 July 2021).

Follesdal, A. (2005) 'Human Rights and Relativism', in Follesdal, A. and Pogge, T. (eds) *Real World Justice: Grounds, Principles, Human Rights, and Social Institutions*. Dordrecht:

Springer Netherlands (Studies in Global Justice), pp. 265–283. doi:10.1007/1-4020-3142-4_15.

Fontana, B. (2008) 'Hegemony and power in Gramsci', in Howson, R. and Smith, K. (eds) *Hegemony: Studies in Consensus and Coercion*. New York and Abingdon: Routledge, pp. 92–118.

Foucault, M. (1977) *Discipline and punish: The Birth of the Prison*. Translated by Alan Sheridan. London: Allen Lane.

Fourcade, M. and Gordon, J. (2020) 'Learning Like a State: Statecraft in the Digital Age', *Journal of Law and Political Economy*, 1(1). doi:10.5070/LP61150258.

Fourcade, M. and Healy, K. (2017) 'Seeing like a market', *Socio-Economic Review*, 15(1), pp. 9–29. doi:10.1093/ser/mww033.

Fraser, N. (2008) 'Abnormal Justice', *Critical Inquiry*, 34(3), pp. 393–422. doi:10.1086/589478.

Fraser, N. (2010) 'Who Counts? Dilemmas of Justice in a Postwestphalian World', *Antipode*, 41, pp. 281–297. doi:10.1111/j.1467-8330.2009.00726.x.

Fusch, P. and Ness, L.R. (2015) 'Are We There Yet? Data Saturation in Qualitative Research.', *The Qualitative Report*, 20(9), pp. 1408–1416.

Fussey, P. and Murray, D. (2019) *Independent report on the London Metropolitan Police Service's trial of live facial recognition technology.* UK: The Human Rights Big Data and Technology Project. Available at: https://repository.essex.ac.uk/24946/1/London-Met-Police-Trial-of-Facial-Recognition-Tech-Report-2.pdf (Accessed: 6 April 2021).

Galič, M., Timan, T. and Koops, B.-J. (2017) 'Bentham, Deleuze and Beyond: An Overview of Surveillance Theories from the Panopticon to Participation', *Philosophy & Technology*, 30(1), pp. 9–37. doi:10.1007/s13347-016-0219-1.

Gandy Jr, O.H. (1993) *The Panoptic Sort: A Political Economy of Personal Information.* Boulder: Westview Press.

Gandy Jr, O.H. (2021) *The Panoptic Sort: A Political Economy of Personal Information*. 2nd edn. New York: Oxford University Press.

Garland, D. (2001) *The Culture of Control: Crime and Social Order in Contemporary Society*. Chicago: The University of Chicago Press.

Garland, D. (2004) 'Beyond the culture of control', *Critical review of international social and political philosophy*, 7(2), pp. 160–189.

Gates, K.A. (2011) *Our biometric future: Facial recognition technology and the culture of surveillance*. New York and London: New York University Press.

Gemeente Amsterdam (2019) *Top 600*. Available at: https://www.amsterdam.nl/wonen-leefomgeving/veiligheid/top600/ (Accessed: 3 February 2019).

Gemeente Den Haag (2020) *Landelijke evaluatie MEOS app*. Available at: https://denhaag.notubiz.nl/modules/13/overige_bestuurlijke_stukken/566779?parent_event=650632 (Accessed: 4 October 2021).

Gibson, J.J. (1979) 'The theory of affordances', in Gieseking, J.J. et al. (eds) *The People, Place and Space Reader*. New York and Abingdon: Routledge; Taylor & Francis Group, pp. 67–82.

Giddens, A. (1984) *The constitution of society: Outline of the theory of structuration*. Berkeley and Los Angeles: University of California Press.

Golumbia, D. (2009) *The Cultural Logic of Computation*. Cambridge, MA and London, England: Harvard University Press.

Gorcsosova, E. (2016) 'The National Law Enforcement Data Programme (NLEDP) Suppliers' Engagement Event', *Partnership for Conflict, Crime & Security Research*, 7 April. Available at: https://www.paccsresearch.org.uk/event/nledp/ (Accessed: 12 July 2021).

Government of the Netherlands (2011) *Police - Government.nl*. Ministerie van Algemene Zaken. Available at: https://www.government.nl/topics/police (Accessed: 3 October 2021).

Grimmelikhuijsen, S. (2012) 'A good man but a bad wizard. About the limits and future of transparency of democratic governments', *Information Polity*, 17(3, 4), pp. 293–302.

Guest, G., Bunce, A. and Johnson, L. (2006) 'How many interviews are enough? An experiment with data saturation and variability', *Field Methods*, 18(1), pp. 59–82.

Haggerty, K.D. (2006) 'Tear down the walls: on demolishing the panopticon', in Lyon, D. (ed.) *Theorizing surveillance; The Panopticon and beyond*. Devon and Oregon: Willan Publishing, pp. 37–59.

Haggerty, K.D. and Ericson, R.V. (2000) 'The surveillant assemblage', *The British journal of sociology*, 51(4), pp. 605–622.

Haggerty, K.D., Wilson, D. and Smith, G.J. (2011) 'Theorizing surveillance in crime control', *Theoretical criminology*, 15(3), pp. 231–237.

Hamilton, A.M. (2020) 'A Genealogy of Critical Race and Digital Studies: Past, Present, and Future', *Sociology of Race and Ethnicity*, 6(3), pp. 1–10.

Hanks, C. (ed.) (2010) *Technology and Values: Essential Readings*. Malden MA, Oxford, West Sussex: Wiley-Blackwell.

Harari, Y.N. (2017) *Homo Deus: A Brief History of Tomorrow*. New York: HarperCollins.

Harcourt, B.E. (2008) *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.

Hardy, J. (2014) *Critical Political Economy of the Media: An Introduction*. New York and London: Routledge.

Hardyns, W. and Rummens, A. (2018) ‘Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges’, *European Journal on Criminal Policy and Research*, 24(3), pp. 201–218. doi:10.1007/s10610-017-9361-2.

Harvey, D. (2005) *The new imperialism*. Oxford: Oxford University Press.

Harvey, D. (2007) *The Limits to Capital*. 3rd edn. London and New York: Verso.

Heeks, R. and Renken, J. (2018) ‘Data justice for development: What would it mean?’, *Information Development*, 34(1), pp. 90–102.

Heidegger, M. (2010) ‘The question concerning technology and other Essays’, in Hanks, C. (ed.) *Technology and Values; Essential Readings*. Malden MA, Oxford, West Sussex: Wiley-Blackwell.

Hendrix, J.A. *et al.* (2018) ‘The Eyes of Law Enforcement in the New Panopticon’:, *Surveillance & Society*, 16(1), pp. 53–68.

Henman, P. (2010) *Governing electronically: E-government and the reconfiguration of public administration, policy and power*. Basingstoke and New York: Palgrave Macmillan.

Henman, P. (2011) ‘Conditional citizenship? Electronic networks and the new conditionality in public policy’, *Policy & Internet*, 3(3), pp. 1–18.

HLN (2016) ‘Politie wordt iPolice: intelligent en digitaal’, *HLN*, 20 May. Available at: https://www.hln.be/nieuws/binnenland/politie-wordt-ipolice-intelligent-en-digitaal~acf1dcf2/ (Accessed: 5 September 2019).

Hobolt, S.B. (2018) ‘The crisis of legitimacy of European institutions’, in Castells, M. et al. (eds) *Europe’s Crises*. Cambridge UK; Malden, MA: Polity Press, pp. 243–268.

Hoffmann, A.L. (2017) ‘Beyond distributions and primary goods: Assessing applications of Rawls in information science and technology literature since 1990’, *Journal of the Association for Information Science and Technology*, 68(7), pp. 1601–1618. doi:10.1002/asi.23747.

Hogan, L. (2011) ‘Human Rights and the Politics of Universality’, *Louvain Studies*, 35(1–2), pp. 181–199.

Home Office (2016) *Home Office (HO) National Law Enforcement Data Programme (NLEDP) Application Development Service - Digital Marketplace*, GOV.UK Digital Marketplace. Available at: https://www.digitalmarketplace.service.gov.uk/digital-outcomes-and-specialists/opportunities/1227 (Accessed: 1 December 2021).

Home Office (2018) *Transparency data Police transformation fund: investments in 2018 to 2019.* GOV.UK. Available at: https://www.gov.uk/government/publications/police-transformation-fund-investments-in-2018-to-2019 (Accessed: 20 July 2018)

Hutchby, I. (2001) ‘Technologies, texts and affordances’, *Sociology*, 35(2), pp. 441–456.

ICO (2018) *ICO finds Metropolitan Police Service’s Gangs Matrix breached data protection laws*, *Information Commissioner’s office*. Available at: https://ico.org.uk/about-the-ico/news-and-

events/news-and-blogs/2018/11/ico-finds-metropolitan-police-service-s-gangs-matrix-breached-data-protection-laws/ (Accessed: 1 December 2021).

Inspectie Openbare Orde en Veiligheid (2008) *Informatiegestuurde Politie*. Den Haag: Ministerie van Binnenlandse Zaken en Koningsrelaties. Available at: https://adoc.pub/informatiegestuurde-politie.html (Accessed: 30 November 2021).

Interpol (2018) *Speaker 04*. Available at: https://www.youtube.com/watch?v=foXSJCtHSqs (Accessed: 18 June 2021).

Introna, L. and Wood, D. (2004) 'Picturing algorithmic surveillance: the politics of facial recognition systems', *Surveillance & Society*, 2(2/3), pp. 177–198.

Isaac, W. and Lum, K. (2018) 'Setting the Record Straight on Predictive Policing and Race', *In Justice Today*, 3 January. Available at: https://medium.com/in-justice-today/setting-the-record-straight-on-predictive-policing-and-race-fe588b457ca2 (Accessed: 6 September 2021).

Jackson, J. *et al.* (2012) *Just authority? Public trust and police legitimacy*. London: Willan.

Jackson, J. and Bradford, B. (2009) 'Crime, policing and social order: on the expressive nature of public confidence in policing', *The British Journal of Sociology*, 60(3), pp. 493–521.

James, J. (2005) *The New Abolitionists: (Neo) Slave Narratives and Contemporary Prison Writings*. Albany: State University of New York Press.

Jansen, F. (2017) 'Smart Borders: Challenges and Limitations of Data-Driven Borders.', in Grzinic, M. (ed.) *Border Thinking: Disassemble Histories of racial violence.* Berlin: Sternberg Press.

Jansen, F. (2018) *Data Driven Policing in the Context of Europe*. Data Justice Lab, pp. 1–17. Available at: https://datajusticeproject.net/wp-content/uploads/sites/30/2019/05/Report-Data-Driven-Policing-EU.pdf.

Jansen, F. (2020) 'Explaining Society: Critical Realism in the Social Sciences', *European Journal of Communication*, 35(3), pp. 308–310.

Jansen, F. (2021) 'The State: A Key Actor in Shaping Data Infrastructure Space', in Herlo, B. et al. (eds) *Practicing Sovereignty: Digital Involvement in Times of Crises*. Bieleveld: Transaction Verlag, pp. 133–146. Available at: https://www.transcript-publishing.com/media/pdf/66/b3/2a/oa9783839457603.pdf.

Jansen, F., Sánchez-Monedero, J. and Dencik, L. (2021) 'Biometric identity systems in law enforcement and the politics of (voice) recognition: The case of SiiP', *Big Data & Society*, 8(2), pp. 1–13. doi:10.1177/20539517211063604.

Jansen, F. (forthcoming) 'Predictive policing: transforming the city into a medium for control', in Currie, M. Know, J. McGregor, C. (eds) *Data Justice and the Right to the City.* Edinburgh: Edinburgh University Press.

Jeppesen, S. and Sartoretto, P. (eds) (2020) *Media Activist Research Ethics: Global Approaches to Negotiating Power in Social Justice Research*. Cham: Springer Nature.

Jessop, B. (2001) ‘Institutional Re(turns) and the Strategic–Relational approach’, *Environment and planning A: Economy and Space*, 33(7), pp. 1213–1235. doi:10.1068/a32183.

Jessop, B. (2005) ‘Critical Realism and the Strategic-Relational Approach’, *New Formations*, 56(1), pp. 40–53.

Jessop, B. (2012) ‘The state’, in Fine, B., Saad-Filho, A., and Boffo, M. (eds) *The Elgar Companion to Marxist Economics*. Cheltenham and Northampton: Edward Elgar Publishing, pp. 333–340.

Jessop, B. (2016) *The State: Past, Present, Future*. Cambridge UK; Malden, MA: Polity Press.

Johnson, J. (2016) ‘The question of information justice’, *Communications of the ACM*, 59(3), pp. 27–29. doi:10.1145/2879878.

Johnson, J.A. (2014) ‘From open data to information justice’, *Ethics and Information Technology*, 16(4), pp. 263–274. doi:10.1007/s10676-014-9351-8.

Johnston, R.G. and Warner, J.S. (2010) ‘Security Theater in Future Arms Control Regimes’, in *Proceedings of the 51st INMM Meeting*. Baltimore, Maryland, USA, pp. 11–15.

Jones, C. (2011) *Implementing the “principle of availability”: The European Criminal Records Information System The European Police Records Index System The Information Exchange Platform for Law Enforcement Authorities.* Statewatch. Available at: http://www.statewatch.org/analyses/no-145-ecris-epris-ixp.pdf (Accessed: 28 November 2021).

Jones, C. (2017) *Market Forces: the development of the EU security-industrial complex,* TNI and Statewatch. Available at: https://www.statewatch.org/publications/reports-and-books/market-forces-the-development-of-the-eu-security-industrial-complex/ (Accessed: 28 November 2021).

Kak, A. (2020) ‘Introduction’, in Kak, A. (ed.) *Regulating Biometrics: Global Approaches and Open Questions,* AI Now Institute, pp. 62–69. Available at: https://ainowinstitute.org/regulatingbiometrics.pdf.

Katzenbach, C. and Ulbricht, L. (2019) ‘Algorithmic governance’, *Internet Policy Review*, 8(4), pp. 1–18. doi:10.14763/2019.4.1424.

Kaufmann, M., Egbert, S. and Leese, M. (2019) ‘Predictive Policing and the Politics of Patterns’, *The British Journal of Criminology*, 59(3), pp. 674–692. doi:10.1093/bjc/azy060.

Kayser-Bril, N. (2020) *At least 11 police forces use face recognition in the EU, AlgorithmWatch reveals,* Algorithm Watch. Available at: https://algorithmwatch.org/en/face-recognition-police-europe/ (Accessed: 3 November 2021).

Kilkenny, M.F. and Robinson, K.M. (2018) ‘Data quality: “Garbage in–garbage out”’, *Health Information Management Journal*, 47(3), pp. 103–105. doi:10.1177/1833358318774357.

Kind, C. (2019) *Biometrics and facial recognition technology – where next?,* Ada Lovelace Institute. Available at: https://www.adalovelaceinstitute.org/blog/biometrics-and-facial-recognition-technology-where-next/ (Accessed: 6 April 2021).

Kindt, E. (2020) ‘A First Attempt at Regulating Biometric Data in the European Union.’, in Kak, A. (ed.) *Regulating Biometrics: Global Approaches and Open Questions*. AI Now Institute, pp. 62–69. Available at: https://ainowinstitute.org/announcements/ai-now-launches-regulating-biometrics-global-approaches-and-open-questions.html.

Kofman, A.K. (2018) ‘Interpol Rolls Out International Voice Identification Database Using Samples From 192 Law Enforcement Agencies’, *The Intercept*, 25 June. Available at: https://theintercept.com/2018/06/25/interpol-voice-identification-database/ (Accessed: 6 April 2021).

Koning, B. de (2015) ‘Waarom moest de Nationale Politie er ook alweer komen?’, *De Correspondent*, 2 June. Available at: https://decorrespondent.nl/2891/waarom-moest-de-nationale-politie-er-ook-alweer-komen/395570667338-068b78c7 (Accessed: 1 October 2021).

Korpsleiding (2021) *Besluit wob verzoek criminaliteits anticipatie systeem geredigeerd,* Nationale politie. Available at: https://www.politie.nl/binaries/content/assets/politie/wob/00-landelijk/criminaliteits-anticipatie-systeem/20210406---8313---besluit-wob-verzoek-criminaliteits-anticipatie-systeem_geredigeerd.pdf (Accessed: 1 December 2021).

Kressel, K. (1985) *The process of divorce: How professionals and couples negotiate settlements*. New York: Basic Books.

Langlois, G. and Elmer, G. (2019) ‘Impersonal subjectivation from platforms to infrastructures’, *Media, Culture & Society*, 41(2), pp. 236–251. doi:10.1177/0163443718818374.

Lee, M. and McGovern, A. (2013) ‘Force to sell: Policing the image and manufacturing public confidence’, *Policing and Society*, 23(2), pp. 103–124.

Leese, M. (2020) ‘Fixing State Vision: Interoperability, Biometrics, and Identity Management in the EU’, *Geopolitics*, pp. 1–21. doi:10.1080/14650045.2020.1830764.

Lemke, T. (2002) ‘Foucault, governmentality, and critique’, *Rethinking Marxism*, 14(3), pp. 49–64. doi:10.1080/089356902101242288.

Leufer, D. and Jansen, F. (2020) ‘The EU is funding dystopian Artificial Intelligence projects’, *Euractiv*, 22 January. Available at: https://www.euractiv.com/section/digital/opinion/the-eu-is-funding-dystopian-artificial-intelligence-projects/ (Accessed: 27 August 2021).

Liberty (2018) *Cardiff resident launches first UK legal challenge to police use of facial recognition technology in public spaces*. Available at: https://www.libertyhumanrights.org.uk/news/press-releases-and-statements/cardiff-resident-launches-first-uk-legal-challenge-police-use (Accessed: 22 June 2019).

Loader, I. (1997) ‘Policing and the social: Questions of symbolic power’, *British Journal of Sociology*, 48(1), pp. 1–18.

Loader, I. and Mulcahy, A. (2001) ‘The power of legitimate naming: Part I—chief constables as social commentators in post-war England’, *British Journal of Criminology*, 41(1), pp. 41–55.

London Policing Ethics Panel (2018) *London Policing Ethics Panel: Interim report on Live Facial Recognition*. London: Mayor of London. Available at: http://www.policingethicspanel.london/uploads/4/4/0/7/44076193/lpep_report_-_live_facial_recognition.pdf (Accessed: 4 October 2021).

Lum, K. and Isaac, W. (2016) 'To predict and serve? In: Significance.', *Significance*, 13(5), pp. 14–19. doi:10.1111/j.1740-9713.2016.00960.x.

Lyon, D. (2001) *Surveillance society: Monitoring everyday life*. Buckingham: Open University Press.

Lyon, D. (2003) 'Surveillance as social sorting; Computer codes and mobile bodies', in Lyon, D. (ed.) *Surveillance as Social Sorting: Privacy, risk, and digital discrimination*. London and New York: Routledge.

Lyon, D. (2006a) 'The search for surveillance theory', in Lyon, D. (ed.) *Theorizing Surveillance; The panopticon and beyond*. Devon and Oregon: Willan Publishing.

Lyon, D. (ed.) (2006b) *Theorizing surveillance; The panopticon and beyond*. Devon and Oregon: Willan Publishing.

Lyon, D. (2007) *Surveillance studies: An overview*. Cambridge UK; Malden, MA: Polity Press.

Lyon, D. (2008) 'Biometrics, identification and surveillance', *Bioethics*, 22(9), pp. 499–508. doi:10.1111/j.1467-8519.2008.00697.x.

Martin, A.K., Van Brakel, R.E. and Bernhard, D.J. (2009) 'Understanding resistance to digital surveillance: Towards a multi-disciplinary, multi-actor framework', *Surveillance & Society*, 6(3), pp. 213–232.

Martin, R. and Bradford, B. (2021) 'The anatomy of police legitimacy: Dialogue, power and procedural justice', *Theoretical Criminology*, 25(4), pp. 559–577. doi:10.1177/1362480619890605.

Mayer-Schönberger, V. and Cukier, K. (2013) *Big data: a revolution that will transform how we live, work and think*. London: Murray.

Mayor of London (2018) *Review of the MPS Gangs Matrix*, *London City Hall*. Available at: https://www.london.gov.uk//mopac-publications-0/review-mps-gangs-matrix (Accessed: 1 December 2021).

Mazzucato, M. (2011) 'The entrepreneurial state', *Soundings*, 49(49), pp. 131–142.

McEwen, C.A. and Maiman, R.J. (1981) 'Small claims mediation in Maine: An empirical assessment', in Menkel-Meadow, C. (ed.) *Mediation; Theory, Policy and Practice*. 1st edn. New York and Abingdon: Routledge, p. 237.

McGee, R. (2010) *Synthesis Report: Review of Impact and Effectiveness of Transparency and Accountability Initiatives*. SSRN Scholarly Paper ID 2188139. Rochester, NY: Social Science Research Network. doi:10.2139/ssrn.2188139.

McKechnie, A. (2007) *Beyond Barriers: A critical Realist perspective on disability and the meaning of the dwelling*. PhD Thesis. Cardiff University.

McQuillan, D. (2015) 'Algorithmic states of exception', *European Journal of Cultural Studies*, 18(4–5), pp. 564–576. doi:10.1177/1367549415577389.

Melgaço, L. and Van Brakel, R. van (2021) 'Smart Cities as Surveillance Theatre', *Surveillance & Society*, 19(2), pp. 244–249. doi:10.24908/ss.v19i2.14321.

Metcalfe, P. and Dencik, L. (2019) 'The politics of big borders: Data (in) justice and the governance of refugees', *First Monday*, 24(4). doi:https://doi.org/10.5210/fm.v24i4.9934.

Metcalfe, P. and Jansen, F. (forthcoming) 'The Facade of Datafication; The Real Harms of an Imagined World', in *Vertical Atlas*.

Miller, P. and Rose, N. (1990) 'Political rationalities and technologies of government', in Hanninen, S. and Palonen, K. (eds) *Texts, contexts, concepts. studies on politics and power in language*. Helsinki: Finnish Political Science Association, pp. 166–183.

Ministerie van AlgemeneZaken (2011) *Investigation and prosecution of criminals*. Available at: https://www.government.nl/topics/crime-and-crime-prevention/investigation-and-prosecution-of-criminals (Accessed: 2 December 2021).

Morrison, G.S. *et al.* (2016) 'INTERPOL survey of the use of speaker identification by law enforcement agencies', *Forensic Science International*, 263, pp. 92–100. doi:10.1016/j.forsciint.2016.03.044.

Mulcahy, A. (2013) *Policing Northern Ireland*. Abingdon and New York: Routledge.

Murdock, J. (2016) 'UK police and Home Office creating centralised intelligence database "without parliamentary oversight"', *International Business Times*, 18 November. Available at: https://www.ibtimes.co.uk/mysterious-uk-surveillance-system-will-store-billions-records-without-parliamentary-oversight-1592332 (Accessed: 12 July 2021).

Nationale politie (2017) *Politie wint ICT-prijs*. Available at: https://www.politie.nl/nieuws/2017/november/1/00-de-politie-wint-belangrijke-ict-prijs.html (Accessed: 1 December 2021).

Nationale politie (2019a) *ANPR*. Available at: https://www.politie.nl/themas/anpr.html#alinea-title-hoe-lang-worden-de-gegevens-bewaard (Accessed: 1 December 2021).

Nationale politie (2019b) *Misdaadcijfers 2018 verder gedaald*. Available at: https://www.politie.nl/nieuws/2019/januari/17/misdaadcijfers-2018-verder-gedaald.html (Accessed: 14 July 2021).

Nationale politie (2020) *Statusupdate high risk applicaties PSbD*. Available at: https://www.bitsoffreedom.nl/wp-content/uploads/2020/10/rapport-statusupdate-high-risk-applicaties-psbd-12-2019-v1.0_def.pdf (Accessed: 30 November 2021).

Nationale politie (2021) *Wat doet de politie tegen cybercrime?* Available at: https://www.politie.nl/informatie/wat-doet-de-politie-tegen-cybercrime.html (Accessed: 2 December 2021).

Newell, B. (2021) ‘Introduction: Surveillance and the COVID-19 Pandemic: Views from Around the World’, *Surveillance & Society*, 19(1), pp. 81–84. doi:10.24908/ss.v19i1.14606.

Niklas, J. and Dencik, L. (2021) ‘What rights matter? Examining the place of social rights in the EU’s artificial intelligence policy debate’, *Internet Policy Review*, 10(3). doi:10.14763/2021.3.1579.

NIST (2020) *Face Recognition Vendor Test (FRVT)*. Available at: https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt (Accessed: 5 April 2021).

NOS nieuws (2017) ‘Fiscus mag beelden snelwegcamera’s niet gebruiken’, *NOS*, 24 February. Available at: https://nos.nl/l/2159788 (Accessed: 1 December 2021).

O’Mahoney, J. and Vincent, S. (2014) ‘Critical Realism as an Empirical Project: A Beginners Guide.’, in Edwards, P.K., O’Mahoney, J., and Vincent, S. (eds) *Studying Organizations Using Critical Realism: A Practical Guide.* Oxford: OUP.

Omi, M. and Winant, H. (1986) *Racial formation in the United States: From the 1960s to the 1980s.* New York: Routledge and Kegan Paul.

Openbaar Ministerie (2019) *Top 600 aanpak - Arrondissementsparket Amsterdam - Openbaar Ministerie*. Available at: https://www.om.nl/organisatie/arrondissementsparket-amsterdam/top-600 (Accessed: 4 October 2021).

Owusu-Bempah, A. (2017) ‘Race and policing in historical context: Dehumanization and the policing of Black people in the 21st century’, *Theoretical Criminology*, 21(1), pp. 23–34.

Pacheco-Vega, R. and Parizeau, K. (2018) ‘Doubly Engaged Ethnography: Opportunities and Challenges When Working With Vulnerable Communities’, *International Journal of Qualitative Methods*, 17(1), pp. 1–17. doi:10.1177/1609406918790653.

Patton, M.Q. (2002) *Qualitative evaluation and research methods*. Thousand Oaks; London; New Delhi: SAGE Publications, inc.

Pearsall, B. (2010) ‘Predictive policing: The future of law enforcement’, *National Institute of Justice Journal*, 266(1), pp. 16–19.

Peña Gangadharan, S. and Niklas, J. (2019) ‘Decentering technology in discourse on discrimination’, *Information, Communication & Society*, 22(7), pp. 882–899. doi:10.1080/1369118X.2019.1593484.

Perry, W.L. *et al.* (2013) *Predictive policing: the role of crime forecasting in law enforcement operations*. Santa Monica, CA: RAND.

Peters, J. (2013) ‘Calendar, clock, tower.’, in Stolow, J. (ed.) *Deus in Machina: Religion and Technology in Historical Perspective.* New York: Fordham University Press, pp. 25–42.

Police.uk (2019) *Automatic Number Plate Recognition*. Available at: https://www.police.uk/information-and-advice/automatic-number-plate-recognition/.

Politics.co.uk (2021) *Police Numbers - All you need to know*. Available at: https://www.politics.co.uk/reference/police-numbers-and-recruitment/ (Accessed: 1 October 2021).

Ponterotto, D. (2016) 'Resisting the male gaze: feminist responses to the "normatization" of the female body in Western culture', *Journal of international women's studies*, 17(1), pp. 133–151.

Poster, M. (1990) 'Foucault and data bases', *Discourse*, 12(2), pp. 110–127.

Privacy First (2018) *Interview met Privacy First over nieuwe wet ANPR.* Available at: https://www.privacyfirst.nl/aandachtsvelden/wetgeving/item/1133-interview-met-privacy-first-over-nieuwe-wet-anpr.html (Accessed: 1 December 2021).

Privacy First (2021) *EenVandaag: 'Politie wil door met fotograferen en opslaan miljoenen kentekens, privacywaakhond spant rechtszaak aan*. Available at: https://www.privacyfirst.nl/rechtszaken-1/itemlist/tag/ANPR.html (Accessed: 1 December 2021).

Raley, R. (2013) 'Dataveillance and countervailance', in Gitelman, L. (ed.) *Raw Data is an Oxymoron*. Cambridge MA, London: MIT Press, pp. 121–146.

Rawls, J. (1999) *A theory of justice: Revised edition*. Cambridge, MA: The Belknap Press of Harvard University Press.

Raz, J. (2009) *Between authority and interpretation: On the theory of law and practical reason*. Oxford: OUP.

Reclaim Your Face (2021) *Reclaim Your Face: Ban Biometric Mass Surveillance!*, *Reclaim Your Face*. Available at: https://reclaimyourface.eu/ (Accessed: 25 June 2021).

Redden, J. (2018) *The Harm That Data Do, Scientific American*. Available at: https://www.scientificamerican.com/article/the-harm-that-data-do/ (Accessed: 3 March 2021).

Richard Vis, X. (2020) 'Police use of fingerprint scanners disproportionately targets Black Britons', *Wired UK*, 11 March. Available at: https://www.wired.co.uk/article/police-fingerprint-scan-uk (Accessed: 4 October 2021).

Rijksoverheid (2010) *Organisatie politie - Politie - Rijksoverheid.nl*. Available at: https://www.rijksoverheid.nl/onderwerpen/politie/organisatie-politie (Accessed: 3 October 2021).

Rijksoverheid (2021) *Straffen en maatregelen voor jongeren*. Available at: https://www.rijksoverheid.nl/onderwerpen/straffen-en-maatregelen/straffen-en-maatregelen-voor-jongeren (Accessed: 9 July 2021).

Robinson, S. (2013) 'The interview: A process of qualitative inquiry', in Ellison, E. (ed.) *The International Encyclopedia of Media Studies*. Brisbane: M/C Reviews, pp. 135–154.

Romano, S. (2017) *Moralising Poverty: The 'Undeserving'Poor in the Public Gaze*. London: Routledge.

Roosen, M. (2020) ‘What SyRI Can Teach Us about Technical Solutions for Societal Challenges’, *Global Data Justice*, 20 February. Available at: http://globaldatajustice.org/2020-02-20-roosen-syri/ (Accessed: 8 August 2021).

Rothstein, H.R. and Hopewell, S. (2009) ‘Grey literature’, in Cooper, H., Hedges, L.V., and Valentine, J. (eds) *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation, pp. 103–125.

Sadowski, J. (2019) ‘When data is capital: Datafication, accumulation, and extraction’, *Big Data & Society*, 6(1). doi:10.1177/2053951718820549.

Sayer, A. (1999) *Realism and social science*. London, Thousand Oaks, New Delhi: Sage Publications.

Sayer, A. (2010) *Method in social science: a realist approach*. Rev. 2nd ed. New York and London: Routledge.

Schneier, B. (2003) *Beyond fear: Thinking sensibly about security in an uncertain world.* New York, NY: Copernicus Books.

Scott, J.C. (1998) *Seeing like a state: How certain Schemes to Improve the Human Condition Have Failed.* New York and London: Yale University Press.

Scott, S. (2018) *The War on Gangs or a Racialised War on Working Class Black Youths.* London: The Monitoring Group.

Seale, C. (1998) ‘Qualitative interviewing’, in Seale, C. (ed.) *Researching society and culture*. London: Sage Publications, pp. 202–216.

Selbst, A.D. *et al.* (2019) ‘Fairness and abstraction in sociotechnical systems’, in *Proceedings of the conference on fairness, accountability, and transparency*. *FAT*19*, New York, NY, USA: Association for Computing Machinery, pp. 59–68. doi:DOI:https://doi.org/10.1145/3287560.3287598.

Sen, A. (2005) ‘Human Rights and Capabilities’, *Journal of Human Development*, 6(2), pp. 151–166. doi:10.1080/14649880500120491.

Sen, A.K. (2009) *The idea of justice*. Cambridge, MA and London, England: Harvard University Press.

Shapiro, A. (2017) ‘Reform predictive policing’, *Nature*, 541, pp. 458–460. doi:10.1038/541458a.

Silverstone, R., Hirsch, E. and Morley, D. (1991) ‘Listening to a long conversation: an ethnographic approach to the study of information and communication technologies in the home’, *Cultural Studies*, 5(2), pp. 204–227.

Smith, C. and Elger, T. (2014) ‘Critical Realism and interviewing Subjects’, in Edwards, P.K., O’Mahoney, J., and Vincent, S. (eds) *Studying Organizations Using Critical Realism: A Practical Guide*. Oxford: OUP.

South Wales Police (2021) *Facial recognition deployments 2017-2020*. Available at: https://www.south-wales.police.uk/SysSiteAssets/media/downloads/south-wales/about-us/frt/FRT-deployments.pdf (Accessed: 30 November 2021).

Srnicek, N. (2017) *Platform Capitalism*. Cambridge UK and Malden, MA: Polity Press.

Star, S.L. (1999) 'The ethnography of infrastructure', *American Behavioral Scientist*, 43(3), pp. 377–391.

Statt, N. (2020) 'Controversial facial recognition firm Clearview AI facing legal claims after damning NYT report', *The Verge*. Available at: https://www.theverge.com/2020/1/24/21079354/clearview-ai-nypd-terrorism-suspect-false-claims-facial-recognition (Accessed: 27 January 2021).

Sunshine, J. and Tyler, T.R. (2003) 'The role of procedural justice and legitimacy in shaping public support for policing', *Law & Society Review*, 37(3), pp. 513–548.

Surveillance Camera Commissioner (2016) *Annual Report 2015/16*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/569559/57586_unnum_camera_WEB.PDF.

Tankebe, J. (2014a) 'Police legitimacy', in Reisig, M. and Kane, R. (eds) *The Oxford handbook of police and policing*. New York: Oxford University Press, pp. 238–259.

Tankebe, J. (2014b) 'Rightful authority: Exploring the structure of police self-legitimacy', *Available at SSRN 2499717* [Preprint]. doi:https://ssrn.com/abstract=2499717 or http://dx.doi.org/10.2139/ssrn.2499717.

Tate, S.A. (2016) '"I can't quite put my finger on it": Racism's touch', *Ethnicities*, 16(1), pp. 68–85.

Taylor, L. (2017) 'What is data justice? The case for connecting digital rights and freedoms globally', *Big Data & Society*, 4(2), pp. 1–14. doi:10.1177/2053951717736335.

Taylor, L. *et al.* (eds) (2021) *Data justice and COVID-19*. London: Meatspace Press. Available at: https://meatspacepress.com/go/data-justice-and-covid-19-internet-archive/ (Accessed: 29 November 2021).

Taylor, L. and Broeders, D. (2015) 'In the name of Development: Power, profit and the datafication of the global South', *Geoforum*, 64, pp. 229–237. doi:10.1016/j.geoforum.2015.07.002.

Thatcher, J., O'Sullivan, D. and Mahmoudi, D. (2016) 'Data colonialism through accumulation by dispossession: New metaphors for daily data', *Environment and Planning D: Society and Space*, 34(6), pp. 990–1006. doi:10.1177/0263775816633195.

Tongco, M.D.C. (2007) 'Purposive sampling as a tool for informant selection', *Ethnobotany Research and applications*, 5, pp. 147–158.

Tuckett, A.G. (2005) 'Part II. Rigour in qualitative research: complexities and solutions', *Nurse researcher*, 13(1), pp. 29–42.

Tyler, T. (2006) *Why People Obey the Law*. New Haven, CT: Yale University Press.

Tyler, T. and Jackson, J. (2013) ‘Future challenges in the study of legitimacy and criminal justice’, *Yale Law School, Public Law Working Paper 264*. Available at: SSRN: https://ssrn.com/abstract=2141322 or http://dx.doi.org/10.2139/ssrn.2141322.

Tyler, T.R. (2003) ‘Procedural justice, legitimacy, and the effective rule of law’, *Crime and Justice; A Review Of Research*, 30, pp. 283–357. doi:10.1086/652233.

Tyler, T.R. (2007) *Legitimacy and Criminal Justice: An International Perspective*. New York: The Russell Sage Foundation.

Van Brakel, R. (2016) ‘Pre-Emptive Big Data Surveillance and its (Dis)Empowering Consequences: The Case of Predictive Policing’, in Van der Sloot, B., Broeders, D., and Schrijvers, E. (eds) *Exploring the Boundaries of Big Data*. Amsterdam: Amsterdam University Press, pp. 117–141. Available at: https://www.ssrn.com/abstract=2772469 (Accessed: 2 March 2021).

Van Brakel, R. (2020a) ‘Een reflectie over het huidig toezicht van het gebruik van surveilancetechnologie door de lokale politie in België’, *Cahiers Politiestudies*, 55, pp. 139–160.

Van Brakel, R. (2020b) ‘Rethinking predictive policing: Towards a holistic framework of democratic algorithmic surveillance.’, in Schuilenburg, M. and Peeters, R. (eds) *The Algorithmic Society Technology, Power, and Knowledge*. Abingdon and New York: Routledge.

Van Brakel, R. (2021a) ‘How to Watch the Watchers? Democratic Oversight of Algorithmic Police Surveillance in Belgium’, *Surveillance & Society*, 19(2), pp. 228–240. doi:10.24908/ss.v19i2.14325.

Van Brakel, R. (2021b) ‘Slaapwandelend stappen we controlemaatschappij in: beschermen we iedereen met de huidige technologie?’, *vrtnws.be*. VRT NEWS, 26 March. Available at: https://www.vrt.be/vrtnws/nl/2021/03/25/opinie-privacy/ (Accessed: 30 March 2021).

Van der Put, C. *et al.* (2013) *Effectief vroegtijdig ingrijpen. Een verkennend onderzoek naar effectief vroegtijdig ingrijpen ter voorkoming van ernstig delinquent gedrag*. Amsterdam: Ministerie van Veiligheid en Justitie., pp. 1–170. Available at: https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/88389.pdf (Accessed: 19 March 2021).

Van Dijck, J. (2014) ‘Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology’, *Surveillance & Society*, 12(2), pp. 197–208. doi:10.24908/ss.v12i2.4776.

Van Dijck, J., Nieborg, D. and Poell, T. (2019) ‘Reframing platform power’, *Internet Policy Review*, 8(2), pp. 1–18.

Van Schendel, S. (2019) ‘Risk Profiling by Law Enforcement Agencies in the Big Data Era: Is There a Need for Transparency?’, in Kosta, E. et al. (eds) *Privacy and Identity Management. Fairness, Accountability, and Transparency in the Age of Big Data: 13th IFIP WG 9.2, 9.6/11.7, 11.6/SIG 9.2.2 International Summer School, Vienna, Austria, August 20-24, 2018, Revised Selected Papers*. Cham: Springer International Publishing (IFIP Advances in Information and Communication Technology), pp. 275–289. doi:10.1007/978-3-030-16744-8_18.

Van Zoonen, L. (2013) ‘From identity to identification: fixating the fragmented self’, *Media, Culture & Society*, 35(1), pp. 44–51.

Varghese, J. (2010) ‘Police structure: A comparative study of policing models’, *Available at SSRN 1605290 or http://dx.doi.org/10.2139/ssrn.1605290* [Preprint].

Vlaanderen (2021) *De organisatie van de politiediensten in België*. Available at: https://www.vlaanderen.be/de-organisatie-van-de-politiediensten-in-belgie (Accessed: 3 October 2021).

Vonk, G. (2014) ‘Repressive Welfare States: The Spiral of Obligations and Sanctions in Social Security’, *European Journal of Social Security*, 16(3), pp. 188–203. doi:10.1177/138826271401600301.

VVSG (2021) *Situering lokale politie*. Available at: https://www.vvsg.be/kennisitem/vvsg/situering-lokale-politie-1 (Accessed: 3 October 2021).

Wacquant, L. (2001) ‘The penalisation of poverty and the rise of neo-liberalism’, *European Journal on Criminal Policy and Research*, 9, pp. 401–412. doi:10.1023/A:1013147404519.

Waddington, D.I. (2005) ‘A Field Guide to Heidegger: Understanding “The Question Concerning Technology”’, *Educational Philosophy and Theory*, 37(4), pp. 567–583. doi:10.1111/j.1469-5812.2005.00141.x.

Wæraas, A. (2009) ‘On Weber: Legitimacy and Legitimation in Public Relations’, in Ihlen, Ø. and Fredriksson, M. (eds) *Public Relations and Social Theory; Key Figures, Concepts and Developments*. 2nd edn. New York and Abingdon: Routledge, pp. 309–330.

Wark, M. (2021) *Capital Is Dead: is this something worse?* London: Verso.

Weber, M. (1968) *Economy and society: An outline of interpretive sociology*. Edited by G. Roth and C. Wittich. Berkeley: University of California Press.

West Midlands Police (2019) *Ethics Committee Briefing Note: DAL_2018_0001_IOM Model*. West Midlands: West Midlands Police and Crime Commissioner. Available at: https://www.westmidlands-pcc.gov.uk/ethics-committee/ethics-committee-reports-and-minutes/.

Wientjes, J. *et al.* (2017) ‘Identifying potential offenders on the basis of police records: development and validation of the ProKid risk assessment tool’, *Journal of Criminological Research, Policy and Practice*, 3(4), pp. 249–260. doi:10.1108/JCRPP-01-2017-0008.

Williams, D. (2014) ‘CAS: Crime Anticipation System: Predictive Policing in Amsterdam’. Available at: https://slidetodoc.com/cas-crime-anticipation-system-predictive-policing-in-amsterdam/ (Accessed: 6 March 2018).

Williams, P. (2015) ‘Criminalising the other: Challenging the race-gang nexus.’, *Race & Class*, 56(3), pp. 18–3.

Williams, P. (2018) *Being Matrixed: the (over)policing of gang suspects in London*. Report. Stopwatch: Research and Actions for Fair and Inclusive Policing. Available at:

http://www.stop-watch.org/news-comment/story/youths-on-gangs-matrix-stopped-and-searched-up-to-three-times-a-day (Accessed: 22 March 2021).

Williams, P. and Clarke, B. (2018) 'The Black Criminal Other as an Object of Social Control', *Social Sciences*, 7(11), p. 234. doi:10.3390/socsci7110234.

Williams, P. and Kind, E. (2019) *Data-Driven Policing: The Hardwiring of Discriminatory Policing Practices Across Europe*. ENAR. Available at: https://www.enar-eu.org/IMG/pdf/data-driven-profiling-web-final.pdf.

Wimmer, J. and Quandt, T. (2007) 'Living in the Risk Society', *Journalism Studies*, 7(2), pp. 336–347. doi:10.1080/14616700600645461.

Young, I.M. (2011) *Justice and the Politics of difference. Princeton University Press.* Princeton NJ: Princeton University Press.

Zachariadis, M., Scott, S. and Barrett, M. (2013) 'Methodological Implications of Critical Realism for Mixed-Methods Research', *MIS Quarterly*, 37(3), pp. 855–879. doi:10.25300/MISQ/2013/37.3.09.

Zenger, R. (2020) 'ICT-systemen politie niet op orde: iedereen de dupe', *Bits of Freedom*, 19 November. Available at: https://www.bitsoffreedom.nl/2020/11/19/ict-systemen-politie-niet-op-orde-iedereen-de-dupe/ (Accessed: 30 November 2021).

Zuberi, T. and Bonilla-Silva, E. (eds) (2008) *White Logic, White Methods: Racism and Methodology*. Lanham, Boulder, New York, Toronto, Plymouth UK: Rowman & Littlefield Publishers.

Zuboff, S. (2015) 'Big other: Surveillance Capitalism and the Prospects of an Information Civilization', *Journal of Information Technology*, 30(1), pp. 75–89. doi:10.1057/jit.2015.5.

Appendix I: research questions for interviews

Expert interviews

- Can you tell me a bit about yourself and the work you are doing on the topic of data and technology?
- Can you describe data-driven policing in your own words?
- Can you give examples of what implementation are happening within your context?
- Are things changing with the introduction of data-driven policing?
- What are the main drivers for these changes?
- What do you consider the benefits of the use of data-driven policing tools?
- What are your concerns with the use of data-driven policing tools?
- Why do you feel police are interested in data-driven policing?
- If we fast forward 5 years, what do you think has changed?
- Is there anything you would like to add or something I haven't asked you?
- Is there anyone else I should talk to?

Police practitioner interviews

- Can you tell me a bit about yourself and how you are involved with this this function [data-driven risk scoring or biometric recognition] ?
- Can you define this function in your own words?
- How is this function being deployed in your police force?
- Why did your department start working on this function?
- Where there any events that triggered the interest in this function?
- How far into the deployment are you?
- What do you believe are he benefits or the opportunities of using a tool like this?
- What challenges have you encountered?
- One of the concerns that people have is that this function is changes the role of the police from more responsive to more pre-emptive interventions. Do you share this concern?
- Another concern is that the use of these functions will lead to the over-policing of certain communities. Do you share this concern?
- How do you ensure that the deployment of these functions are done in the right way?
- Who do you consider your biggest allies and/or critics?
- Is there anything you would like to add or something I haven't asked you?
- Is there anyone else I should talk to?

Civic actors interviews

- Could you briefly tell me a bit about yourself and your work in relation to the police?
- What are the most concerning trends regarding the use of data and technology by police?
- Why do you think police is turning to data and technology?
- What are your specific concerns about this use of data-driven policing technologies?
- Who do you feel are most impacted by its use?
- What do you consider to be the main challenges in working on this topic?
- What do you hope to achieve with your work?
- Who do you consider your key allies?
- Do you engage with police on this matter, yes what have been your experiences and if no, why not?
- There is much critique on police power and their use of data-driven policing tools, what do you want to police to do?
- Is there anything you would like to add or something I haven't asked you?
- Is there anyone else I should talk to?

Appendix II: Information sheet for potential interviewees

Research information

What is this research about?

The aims of the research is to learn about the implications and practices of data-driven decision-making in law enforcement, why police are using these technologies, what are the challenges and opportunities they are confronted with and how this related to the concerns of civil rights groups.

Participation in this project will involve:

Talking about my practices and experiences through an interview. The interview will be digitally recorded for the purposes of the research only. Participation in this project is entirely by choice. Questions are welcome at any time.

What happens with the research information?

All names and identification will be removed or changed in the research so participant contributions are anonymous in any transcriptions or publications and personal identifiable data will be stored separately from the recordings, unless participants have given explicit permission for names, titles, or organisations to be included. The research will be held confidentially at Cardiff University, which is registered with the UK Information Commissioner's Office to process personal data in compliance with Data Protection law and will be stored for four years after the project completion when any personal data will be destroyed.

What are my rights?

You may withdraw from the research at any time without giving a reason and under Data Protection Law may request access to the personal data that is held about you for this study and request that it be deleted. Please contact the PI.

Who is doing this research?

Fieke Jansen is the PhD candidate for this study. The research is conducted under the supervision of Dr Lina Dencik, the Principal Investigator (PI) for this study and a Senior Lecturer at the School of Journalism, Media and Cultural Studies (JOMEC) at Cardiff University. This research is part of the project DATAJUSTICE and is funded by the European Research Council (proposal no. ERC-2017-STG-759903). It has been approved by JOMEC's School Research Ethics Committee.

Researcher contact info:

Fieke Jansen
School of Journalism, Media and Cultural Studies, Cardiff University
Bute Building, King Edward VII Avenue
Cardiff CF10 3NB
United Kingdom
Email: JansenF@cardiff.ac.uk

SCHOOL RESEARCH ETHICS OFFICER CONTACT INFO:

Dr Damien Carney
School of Journalism, Media and Cultural Studies, Cardiff University
Bute Building, King Edward VII Avenue
Cardiff CF10 3NB
United Kingdom
Email: carneyd@cardiff.ac.uk

Where can I find out more about this research and any findings?

Please contact the Principal Investigator, Dr Lina Dencik, for updates. Publications and other resources relating to the research will be available on the project website *https://datajusticeproject.net/*

Appendix III participant observation

- ‘Internet Freedom Festival’. Open Technology Fund. Valencia, 5-8 March 2018.
- ‘Data Sharing for Law Enforcement’. Lorentz Center. Leiden, 31 May 2018.
- ‘Round table on AI and content moderation’. UN Special Rapporteur Freedom of Expression. Geneva, 18 June 2018.
- ‘Anticipating Crime’. Het Nieuwe Instituut. Rotterdam, 11 October 2018.
- ‘Virtual strategy design jam to develop litigating against the use of algorithms in law enforcement in EU’. Digital Freedom Fund. Berlin, 29 January 2019.
- ‘Strategy meeting on strategic litigation human rights and technology in Europe’. Digital Freedom Fund. Berlin, 13 - 14 February 2019.
- ‘Pre-litigating meetings Top 400’. PILP NJCM. Amsterdam, 2019, 2020, 2021.
- ‘Democratizing Data’. Bernstein School of Governance NYU’. New York, 17 - 18 April 2019.
- ‘Expert meeting on Predictive policing. Police and Human Rights Programme’. Amnesty International The Netherlands. Amsterdam, 20 – 21 May 2019.
- ‘Shaping the state of machine learning algorithms within law enforcement’. University of Winchester. London, 6 June 2019.
- ‘Round table meeting ProKid -23’. Dutch National Police. Nieuwegein, 29 October 2019.
- ‘Brainstorm Top 400’. De Brauw Advocated, 1 November 2019.
- ‘The Scottish International Policing Conference’. The University of Edinburgh. Edinburgh, 10 December 2019.
- ‘Hardwiring Discriminatory Police Practices: the Implications of Data-Driven Technological Policing on Minority (Ethnic and Religious) People and Communities’. Open Society Foundation. Barcelona, 30 January 2020.
- ‘Citizen Biometric Panel’. Ada Lovelace. Bristol and online, 1 - 2 February & 9 September 2020.
- ‘City Forum on Policing the Nation’. City Forum. London, 4 - 5th February 2020.
- ‘Chaire Villes et numérique" : Des villes sous surveillance ? Espaces urbains, sécurité et numérique’, Science Po. Paris, 6 February 2020.
- ‘Webinar: Police use of data and tech: what’s the problem?’ OSF Justice Initiative. Online, 4th June 2020.
- ‘Webinar: Police use of data & tech: ways to push back’. OSF Justice Initiative. Online, 18 June 2020.
- ‘Criminal Justice by Algorithm Part I Predictive Policing’. Fair Trials. Online, 13 October 2020.
- ‘Table session: the state of data-driven policing in Europe and its impact on racialised communities’. Justice, Equity, and Technology Table. Online, 15 April 2020.
- ‘Civil Society Discussion on European AI Act – Law Enforcement and high risk’. EDRi. Online, 2019, 2020, 2021.