Doctoral theses at NTNU, 2023:31

Lisa Marie Reutter

Datafication of Public Administration

Between Policy and Practice

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Social and Educational Sciences
Department of Sociology and Political Science

NTNU
Norwegian University of
Science and Technology

Lisa Marie Reutter

Datafication of Public Administration

Between Policy and Practice

Thesis for the Degree of Philosophiae Doctor

Trondheim, February 2023

Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Sociology and Political Science

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Social and Educational Sciences
Department of Sociology and Political Science

ISBN 978-82-326-5905-0 (printed ver.)
ISBN 978-82-326-6686-7 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:31

Printed by NTNU Grafisk senter

Table of Contents

List of Tables

List of Figures

List of papers

This dissertation is based on the following papers, which are referred to in the text by their numbers. They are presented in logical order rather than date of publication.

Paper 1 [published]
Reutter, Lisa 2022. **"Constraining Context: Situating Datafication in Public Administration."** *New Media & Society* 24(4):903-921.

Paper 2 [published]
Broomfield, Heather, and Lisa Reutter 2021. **"Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation."** *Scandinavian Journal of Public Administration* 25(2):73-97.

Paper 3 [published]
Broomfield, Heather, and Lisa Reutter 2022. **"In Search of the Citizen in Public Administration Datafication."** *Big Data & Society* 9(1):1-14.

Paper 4 [Under process at Critical Policy Studies, decision in first round: Revise and Resubmit]
Reutter, Lisa, and Heidrun Åm 2022. **"Public Sector Data as a Resource: Tracing the Emergence and Embedding of a Sociotechnical Imaginary."**

Summary

Algorithms, data platforms, and big data have gained ever more attention in public administration worldwide in the last decade. Policy papers and consultancy reports promise a more effective and seamless public administration that predicts and personalizes citizen needs proactively enabled by these technologies. Simultaneously, social scientists across disciplines have begun to critically investigate this administrative reform, as it has the potential to seriously affect citizen–state relations and increase the public sector's ability to surveil, predict, and classify citizen behavior. The field of critical data and algorithm studies has emerged as a counterforce to rigorously researching the multiple ways data, technology, power, and politics intersect and are integrated into our social lives.

This paper-based dissertation offers unique insight into the inner workings of public administration datafication in Norway and aims to provide a valuable contribution to this emerging body of literature. I define public administration datafication as being made up of two interwoven processes: the use of more and different data and the recirculation of data in increasingly complex ways, both within and outside of the public sector. These intertwined processes are fueled by the idea of data as a resource for improving public administration and society more generally. I research the data-driven public sector both through specific projects where data assemblages are materialized - and a system-level investigation, questioning the emergence and construction of socio-technical imaginaries. I situate datafication beyond the private sector and internet platforms and ask the overall research question of how public administration datafication is framed and embedded in the Norwegian public sector. Norway proves to be an especially interesting case, as the welfare state has collected and stored vast amounts of data for decades, which are now imagined to be recirculated into practice in all aspects of administration.

Through a multi-method research design that includes ethnographic fieldwork among data teams producing technology, interviews with practitioners, a survey, and document analysis, I develop and apply a situated practice approach to datafication. I focus on what practitioners, policy makers, and institutions do, say, and imagine in relation to datafication, rather than measure its immediate effects on society. By drawing attention to silences, unseen issues, and obstacles, I outline new allies for citizen interventions and contestation to avoid determinist accounts of this administrative reform across all four papers. I show that the public sector meets a variety of obstacles when attempting to materialize the sociotechnical imaginary of public sector datafication. In many ways, the data-driven public sector remains a future vision. This then stresses the importance of critically investigating the performativity of sociotechnical imaginaries

that push data-driven technology and the issues omitted and silenced in these future visions. I find that citizens are rarely included in constructing and embedding these sociotechnical imaginaries, as datafication is presented as inevitable, apolitical, and necessary. Contestations and alternative future visions therefore become increasingly difficult.

The four independent papers in this dissertation provide the first empirical and explorative study on public administration datafication in Norway. I invite scholars across disciplines to critically and reflexively question this administrative reform and actively engage in discussions on datafication with policymakers and public sector practitioners.

Acknowledgements

I would like to express my sincere gratitude to the Norwegian public sector, which has welcomed my research ideas with great enthusiasm and provided me with a unique access to its inner workings despite not always agreeing on my critical perspective.

In the last four years, I have again and again imagined myself writing the acknowledgements of my dissertation. However, sitting here, writing the actual text, I am somehow at a loss for words. My PhD journey had its ups and many downs, and I am forever grateful for the support I have received from the people surrounding me over the course of that journey.

First, I want to thank my brilliant partner in crime, Heather Broomfield, for endless discussions on data-driven public administration and other subjects as well. Finding you on the internet was the best thing that has happened to me during my PhD journey. I am honored to have published two amazing papers together with you and look already forward to our next project.

Thank you to my supervisors, Emil Røyrvik and Henrik Spilker, for their valuable input and feedback on this dissertation. A special thanks to Taina Bucher for a fantastic and insightful discussion at my midway seminar. Thank you also to Heidrun Åm for her helpful comments and suggestions on paper drafts and our successful collaboration on the final paper in the thesis.

I was very lucky and began my PhD the same day as my best friend Ingrid Vik Bakken began hers. Thank you for being my best work friend, taking me on writing retreats and coffee and lunch breaks, and encouraging me to finish this thesis. Many of my colleagues have been of great moral support. Special thanks are due to Aurora Sørsveen, Jorunn Kaasa, and Anne Wullum Aasback here.

This dissertation would not have been possible without the unconditional love and support of my friends and family. To Rebecca, Hanne, Henrik, Jens, Katrine, Herman, Anna, Oda, Eirin, Fabian, Anne Line, Mathieu, Miri, Åsta, Ingvill, Sofie, and all my other friends: you have all contributed to the process of completing this dissertation in one way or another through your friendship, phone calls, dinner invitations, concert and theatre attendances, and knitting dates. As I spent the last six months of the writing process in the tiny village of Undredal, I want to thank all the two- and four-legged inhabitants of this village for providing me with the perfect surroundings in which to think and write.

Finally, to my mum and dad, I would like to say the following: thank you for always believing in me and supporting me in whatever crazy idea I happened to be pursuing at the time. Despite not always understanding my work, you encouraged me to find my own way and be curious, never once doubting that I could write a PhD thesis.

The social sciences predominantly achieve external impacts by influencing people to think about things in a different, more precisely reasoned and better informed way, one that will (hopefully) produce better decisions and societal outcomes than would have been achieved without the presence of these disciplines.

(Bastow, Dunleavy, and Tinkler 2014:273)

Introduction to public administration datafication

An interview published on the Norwegian online news platform digi.no in January 2018 begins thus:

> Facebook earns big money with insights based on their users' data. What can Norway learn from Zuckerberg and co? Can we save lives, get people into jobs and education, and create a better society? Yes of course says IT-director Torbjørn Larsen in the Labor and Welfare Administration!

In many ways, this dissertation has its very origin in this quote and what it connotes. Platform companies such as Facebook, Google, and Amazon have based much of their success on the extracting of user data and algorithmic processing of information. However, as this quote illustrates, the idea of data as a resource for value creation and improving society has gained ever more attention beyond private sector business models. A 2016 government white paper, for example, argues that "'It is natural to assume that big data, together with technologies such as automation and artificial intelligence, will be able to change how the public sector delivers and produces its services in the future." (Meld. St. 27 2015-2016: 109). This white paper has since been followed by a variety of strategy and policy papers, in addition to conferences such as the 2019 Building the Country with Data conference, all emphasizing the importance of sharing and using data across the public sector to improve public administration. Data-driven technology is here seen as a savior of the welfare state, which is seemingly threatened by a variety of external forces ranging from an aging population to increased immigration and climate change. The metaphor of "data as the new oil" becomes even more powerful in a country currently in search of something to replace its massive oil sector.

Datafication, the process of quantifying every aspect of life so it can be tabulated, analyzed, and recirculated into practice, presents a profound paradigm shift for public administration and a challenging object of inquiry for social scientists across disciplines. This dissertation aims to critically investigate the inner workings of the data-driven public sector in Norway and explore how datafication is imagined and embedded into practice.

1 Background and motivation

This first chapter positions my own work in a wider landscape of literature on data-driven technology in society and introduces the research aim and purpose of the dissertation. The section thus outlines how the papers presented build on, complement, and challenge existing research in the social sciences.

1.1 Critical data and algorithm studies

We can observe increasing interest in the various ways in which technologies, such as machine learning and big data, impact society among social scientists in recent years (Beer 2017; Bucher 2017; Gillespie 2014a; Kitchin 2014b). Although these technologies have been praised for the endless possibilities they allegedly provide to improve society, a growing body of literature has pointed toward the negative and often unintended effects of data-driven technology on society. This includes both theoretical and empirical accounts of increased commodification, opacity, discrimination, and surveillance and a decrease in control and democratic values in society (Burrell 2016; Eubanks 2018; Lyon 2014; Pasquale 2015; Sadowski 2019). The core of this critique revolves around addressing a change in power dynamics, invoking the idea of a corporate/governmental *inside* in power and control, and a disempowered and unknowing *outside* (Reutter and Spilker 2019). This is also explored in Issar and Aneesh (2022:10) review, who conclude that "concerns about [datafication] can be summed up as the growing institutional capabilities to move contestable issues to a space of reduced negotiability".

The loosely associated and interdisciplinary field of critical data and algorithm studies[1] has emerged as a subfield within the social sciences to critically investigate this paradigm shift and its effect on society (Dalton and Thatcher 2014). Critical data and algorithm studies apply critical social theory "to explore the ways in which [data and algorithms] are never simply neutral objective, independent, raw representations of the world, but are situated, contingent, relational, contextual and do active work in the world" (Kitchin and Lauriault 2014:5), regarding data as produced rather than raw and questioning the positivistic ground on which big data science stands (Gitelman 2013; Iliadis and Russo 2016). There is a variety of conceptual work in the field that has attempted to explain and understand how value-laden data-driven systems are embedded into society and promote specific ways of knowing, as well as exercise power over our social lives (Beer 2017; Boyd and Crawford 2012; Burrell 2016; Gillespie 2014a; Introna 2016; Kitchin 2014a; Klinger and Svensson 2018; Zarsky 2016). Data and technology, how they are conceptualized, who uses them and how, and why they are put to use are central questions asked in this body of literature as datafication is often made to appear invisible and its instinctive power relations obscured by various mechanisms. As Ruppert, Isin, and Bigo (2017:1) highlight, this field "asks questions about the ways in which data has become such an object of power and explores how to critically intervene in its deployment as an object of knowledge" One of the main tasks of critical data and algorithm studies is then to make visible the many ways in which our lives are turned into data and recirculated into action. This dissertation positions itself within critical data and algorithm studies and builds on these central assumptions and questions.

In its first phase, the scholarship associated with critical data and algorithm studies largely focused on private sector datafication efforts, dominated by studies on internet platform

[1] Many of the cited scholars would probably not identify themselves as operating within critical data and algorithm studies. This is solely based on my own perception of this loosely associated and emerging field.

companies, their exploitation of data to make profit, and their use of algorithmic systems to order reality and surveil our everyday lives (see for example Ananny and Crawford 2018; Bucher 2017; Caplan and boyd 2018; Rieder and Sire 2014; Turow and Couldry 2018). Here, I want to highlight Taina Bucher (2012) and Tarleton Gillespie's (2014b) groundbreaking work on platform algorithms, both of which presented my first meeting with this line of thought, instigated my curiosity for studying data and algorithms, in addition to have had a great influence on my own thinking on how to research datafication. Bucher (2012) explored the Facebook algorithm through reverse engineering and showed how social scientists can approach complex technical systems to understand their relevance for society and our social relations. In his 2014 essay on the relevance of algorithms, Gillespie argued that the introduction of these technologies into human knowledge practices has several political ramifications. He called for a sociological inquiry into data-driven technology, which highlights the social processes regarding how technology is made into a legitimate system. These two works also illustrate the close ties between critical data and algorithm and media studies (Flensburg and Lomborg 2021).

Critical studies of data-driven systems have increased exponentially within a short period of time. As datafication enters ever more aspects of social life, empirical research has also begun to extend beyond studies on internet platforms. This includes accounts of the datafication of workplaces (Sánchez-Monedero and Dencik 2019), the role of algorithms in the financial market (Arnoldi 2016; Campbell-Verduyn, Goguen, and Porter 2017), and the datafication of the self, more generally (Cheney-Lippold 2017; Lupton 2016). The arrival of a global pandemic in recent years has once again renewed interest in scholarly and critical accounts of dataveillance, algorithms, and techno-solutionism in the social sciences (Liu and Graham 2021; Milan 2020; Sandvik 2020). Especially relevant for this dissertation, then, is the work done on the datafication of the public sector and the welfare state (Andreassen, Kaun, and Nikunen 2021; Choroszewicz and Mäihäniemi 2020; Dencik 2022; Dencik and Kaun 2020; Nikunen and Hokka 2020; Redden 2018).

1.2 From big tech to public administration

Drawing on Ruppert (2016), among others, Redden (2018: 3) argued that shifting public administration practices due to datafication brings with it a variety of potential implications that social scientists need to be aware of: "these are profound changes in the ways democratic states learn about, engage with and respond to citizens and the information about them" Redden, implications include 1) that citizens become more knowable, traceable across their social lives, and are increasingly transformed into data subjects, 2) that governments are increasingly compelled to share and link data, both across the public sector and externally, 3) that services and decision-making processes are becoming more opaque through automation, 4) that changing power dynamics are making citizens more knowable without citizens having the ability to investigate public sector data practices, and 5) that there is increased use of public-private partnerships in technology development. In other words, datafication has the potential to

radically change public administration and citizen-state relations. Understanding the impact and role of datafication in society becomes especially important when data-driven technology is employed in public administration. As public administration touches upon everybody's life often without any opt-out options, especially in a welfare state context, the public sector's fascination with data and algorithms has therefore tantalized my scholarly curiosity.

First, it is important to recognize that there is a substantial body of literature on digital technology in welfare provision and the public sector in the social sciences (see for example Schou and Hjelholt 2019; Storm Pedersen and Wilkinson 2018). However, data-driven technology brings with it new questions. Emerging empirical research on datafication in public administration includes a variety of case studies on specific public sector projects, such as the integration of data-driven systems in (child) welfare provision (Chaudhuri 2022; Keddell 2015; Redden, Dencik, and Warne 2020), the use of algorithmic systems in unemployment assessment (Allhutter et al. 2020; Zejnilović et al. 2020), and the use of big data and algorithmic decision-making in border regimes (Coulthart and Riccucci 2022) and in the health sector (Hoeyer 2019; Ruckenstein and Schüll 2017). Scholarship on data-driven public administration often focuses on the risks and consequences associated with using data-driven systems in specific contexts and shows how political, social, economic, and cultural factors influence these systems. Interestingly, the idea of predictive policing (Andrejevic, Dencik, and Treré 2020; Benbouzid 2019; Mantello 2016) and smart cities, often linked to public sector applications of data-driven systems, seem to have attracted most critical data and algorithm studies' attention (Cardullo and Kitchin 2019; Große-Bley and Kostka 2021; Hong et al. 2019; Löfgren and Webster 2020; Sadowski and Pasquale 2015). In addition to this focus on applications of datafication, we can observe a slowly emerging literature on public perceptions of algorithmic decision-making and data use (Kennedy et al. 2020, Hintz et al. 2022). This also includes critical accounts of citizens' lack of democratic participation in the making of the data welfare state (van Zoonen 2020). In addition, the imaginaries of public administration datafication drawn up in policy documents have been critically investigated by scholars (Bareis and Katzenbach 2021; Germundsson 2022; Tupasela, Snell, and Tarkkala 2020).

While inquiries into the use of big data and algorithms in public administration are dominated by empirical case studies on specific projects, scholars have also embarked on more conceptual work on the changing relationship between citizens and the state and the understanding of social problems and data justice (Barassi 2019; Dencik et al. 2019; Jørgensen 2021; Kuziemski and Misuraca 2020; Rieder and Simon 2016; Yeung 2017). As Brauneis and Goodman (2018), for example, point out, algorithmic decision-making runs the risk of degrading the decision-making capacity of public servants by creating a great distance between the decisions made and the evidence-gathering that informs these decisions. Central here is also Dencik et al.'s (2019) case study on scoring systems in the UK, in which the authors show how practitioners regard data-driven technology as enabling a "golden view" of society and increasingly engaging in the prediction and risk scoring of citizens. This project is of high relevance to this dissertation, as it

demonstrated the heterogeneity of data systems in use and their contingency on contextual factors and local economic regimes.

As is clearly apparent, most literature on the use of big data and algorithms in public administration has been produced within the last two to three years. Overall, there has been an overwhelming focus on private sector applications of datafication in the first wave of literature associated with critical data and algorithm studies. As datafication gains more ground in public sectors globally, it is key to situate datafication beyond private internet platforms, as demonstrated in this dissertation. Although a body of literature is emerging that investigates the multiple ways data-driven technology is implemented in public administration and impacts citizens, empirical accounts often focus on single fields, such as policing, health, childcare, or education, in isolation. Single-project studies on the issues surrounding data-driven public administration run the risk of disregarding the bigger shifts that the public administration is undergoing and how single projects across the sector interact and are framed. The lack of system level investigations of datafication is therefore addressed in the dissertation. Although datafication is an emerging and relatively new phenomenon, public administration and its practices have long been a key focus of the social sciences. How, then, do public administration studies approach data-driven technology?

1.3 Public administration studies and datafication

There seems to be surprisingly little interest in issues of datafication within the field of public administration studies (Fredriksson et al. 2017). However, this appears to be slowly changing, as AI has attracted some interest in public administration literature more recently (Agarwal 2018; Bullock 2019; Wirtz, Langer, and Fenner 2021). Public administration studies have traditionally paid scant attention to issues regarding the implementation of technology into practice (Meijer 2007), often solely focusing on how to foster technology adoption without discussing its implications for public administration and society in general. These studies argue that public administration must adapt rapidly to the changing technological environment in order to avoid lagging behind (Agarwal 2018; Klievink et al. 2017). In many ways, this body of literature does regard technology as an external force that must be accepted and dealt with, contrary to the social constructivist view of critical data and algorithm studies outlined above.

The main body of literature published within public administration studies has focused on defining this phenomenon and on the possible applications of big data and AI and how to enable change, often praising data-driven technology for its "astonishing positive outcomes for public administration in terms of its efficacy, efficiency, and overall client satisfaction" (Maciejewski 2017:1). Big data and artificial intelligence are seen as major opportunities to improve policy-making, service delivery, and efficiency within public sector management (Chen and Hsieh 2014; Sarker, Wu, and Hossin 2018; Pencheva, Esteve, and Mikhaylov 2020). Research has resulted in a variety of literature on the specific applications of data-driven technology within sectors such as

policing, health, education, and smart cities (Coulthart and Riccucci 2022; Shastri and Deshpande 2020). Papers discussing the challenges of big data and AI in public administration seem to mainly focus on the practical challenges of implementation and management (Agarwal 2018; Fredriksson et al. 2017; Guenduez, Mettler, and Schedler 2020). There is an underlying understanding that the public sector lags behind the private sector and therefore must "become more big data ready in the future" (Klievink et al. 2017:1). Although issues of privacy are referenced in some papers, critical assessments of data-driven technology remain a sidenote in public administration literature. As Andrews (2019), highlights, new technology seems to sweep over public administration, without being critically assessed in the field

This misalignment of focus between critical data and algorithm studies and public administration studies presents an interesting point of departure for this research project. McDonald et al. (2022) encouraged an increased focus of public administration studies on issues of technology and how these might (or might not) be incorporated into the decision-making processes of public administration. At the same time, they also highlight that "Although technology might improve outcomes statistically, it might also pose risks to democratic governance and social equity, among others, that need to be understood" (McDonald et al.:64–65). Indeed, in their recently published literature review on AI in public administration, Wirtz et al. (2021) stressed that the field is struggling to grapple with the societal and political impact of AI on the public sector. This call to engage more deeply with the societal and democratic issues associated with big data and AI in public administration is supported by several other researchers in the field (Andrews 2019; Cordella and Bonina 2012; Veale, Van Kleek, and Binns 2018) and is taken up in this dissertation. Interestingly, Wirtz et al. (2021) also highlighted that citizen perspectives are often absent from research on AI in public administration. This brings us to the last body of literature upon which this dissertation builds: data activism studies.

1.4 Datafication and citizen power

Research associated with critical data and algorithm studies does not only point toward the critical aspects of datafication but has also increasingly focuses on how to mitigate the harm caused by data-driven systems. This then intersects with the rapidly increasing literature on AI and data ethics, examining the ethical aspects of AI, big data, and internet platforms (Cath et al. 2018; Floridi et al. 2018; Mittelstadt et al. 2016; Zwitter 2014). Fairness, accountability, and transparency have become central concepts to discuss and assess the ethical ramifications of data-driven systems across a variety of fields. Although this body of literature has contributed significantly to highlighting the social and cultural aspects of data-driven technology, it is dominated by technical and often isolated approaches to mitigate the impact of AI and big data. There are several calls to move beyond the vague concept of ethics and toward more holistic approaches such as data justice (Dencik, Hintz, and Cable 2016; Taylor 2017). These approaches then mainly investigate power and justice rather than ethics. Both lines of thought do however share a common commitment to challenge data and technology harms.

As Velkova and Kaun (2021:523) pointed out, "most research has focused on negative outcomes, including ethical problems of machine bias and accountability, [and] little has been said about the possibilities of users to resist algorithmic power." This observation has also puzzled me for several years. While there is a growing body of literature pointing toward the social, economic, and political forces shaping datafication, this seemed to have resulted in a focus on negative outcomes and deterministic accounts of datafication's impact on society (a particular deterministic example can here be found in Shosanna Zuboffs work on surveillance capitalism). Once again, this contributes to the idea of a corporate or government "inside" in power and control and a disempowered "outside" increasingly assessed and influenced by data-driven systems. This body of literature often takes for granted that the sociotechnical imaginaries of data-driven public administration materialize in full and therefore also run the risk of overdetermining their power. To avoid determinist accounts of datafication, I have therefore actively highlighted the potential of contestations and resistance across all four papers, rather than solely focusing on negative outcomes and ethical problems of bias, accountability, and transparency. How can datafication be contested beyond ethics?

The most prominent line of thought on data contestation can be found in data activism studies, which address the various ways citizens respond to and resist harmful data practices (Baack 2015; Kennedy 2018; Milan and van der Velden 2016; Lehtiniemi and Ruckenstein 2019). The highly interesting body of literature on data activism draws our analytical attention toward the democratic agency of citizens in studies of datafication and away from notions of citizens as disempowered and unable to challenge their datafied lives (Milan 2017). Data activism studies are sociotechnical, focusing on the political dimension of data and therefore escaping determinism. However, data activism studies have mainly addressed reactive measures taken by citizens, therefore focusing on already existing data systems and how these might be reconfigured by citizens or how publicly available data can be used to improve citizens' everyday lives. We know little about how these systems are framed and produced, which values and agendas steer their development, and the potential of citizen intervention in the early stages of datafication (Flensburg and Lomborg 2021; Hintz et al. 2022). The nascent phase of data-driven public administration and a focus on what happens before data-driven technology is used in specific public sector operations is therefore in focus in this dissertation. Work done by researchers such as McQuillan (2018) or Hintz et al. (2022), focusing on democratic auditing has here been of great inspiration.

1.5 Research aims and purpose

Building upon insights and research from critical data and algorithm studies both on private and public sector application and data activism, in addition to work done in public administration studies, this dissertation aims to critically investigate the inner workings of the data-driven public sector. As with multiple other fields within the social sciences, many of the empirical accounts of

data and technology are situated within a US or UK context. However, public administration differs greatly across countries when it comes to its political framing, scope, size, and function. To produce situated accounts, there is a need for studies in other public administration contexts, which work done by researchers such as Choroszewicz and Mäihäniemi (2020) and Tupasela et al. (2020) clearly demonstrates.

The Nordic public sectors collect, store, and manage enormous amounts of data in order to govern the welfare state, and they are characterized by high democracy scores, which makes them an especially interesting object of investigation (Dencik and Kaun 2020). There has been an overwhelming focus on digitalization and datafication in the Nordic states, stressing the importance of becoming world leaders in public sector technology adaption (Germundsson 2022). Nordic countries keep detailed accounts of their citizens through public registers and often do not provide any method to opt out of this collection (Tupasela et al. 2020). Their public sectors are relatively large and touch upon every aspect of citizens' lives. In addition, the Norwegian public sector in particular enjoys high levels of trust, which often "intersects with a popular belief that technological progress is inevitable, apolitical, and problem free" (Sandvik 2020:2). Datafication has here gained ever more attention and legitimacy in recent years. At the same time, Norway is still in a nascent phase of this development, providing researchers with the opportunity to study the data-driven public sector at an early stage. Public administration is required to provide access to its inner workings to researchers by law, which presents unique possibilities for empirical work on datafication, which has often been hindered by black-boxed processes and industry secrecy (Diakopoulos 2014). Therefore, the Norwegian public sector is especially well-suited to study datafication in the making.

Taking an explorative approach, this research project presents the first comprehensive study on the datafication of public administration in Norway. **Using four papers, the dissertation attempts to answer the overall research question of how datafication is framed and embedded in Norwegian public administration**. Datafication is here understood as a process that aims to change data practices in the public sector, in addition to introducing algorithmic systems into its inner workings. Making use of an abductive research approach as introduced by Tavory and Timmermans (2014) and a multi-method research design, I aim to investigate what practitioners, policy makers, and institutions do and say in relation to datafication, thereby avoiding technocentric accounts of the phenomenon. Zooming in and out, the project offers both insights into changing data practices in specific public sector organizations and system-level accounts of the data-driven public sector. In many ways this is about taking a step back, focusing not on the immediate consequences of datafication but on its construction and embedding in society. By doing so, I aim to find new ways for citizen intervention and configurations. The research project offers a unique and situated empirical account of public administration datafication and contributes significantly to the emerging field of critical data and algorithm

studies, as well as the more established fields of public administration, practice theory, and science and technology studies (STS).

Starting off with a micro-level field study of two data teams in the Norwegian public sector, Paper 1 investigates how the sociotechnical imaginaries of datafication are materialized (or not) in specific data assemblages. This paper focuses on obstacles and challenges in producing data-driven technology in public administration to show how data-driven technology is often constrained in the context of the public sector. While working on Paper 1, I soon realized that the work conducted in the data teams was about more than AI and machine learning. Paper 2, therefore, zooms out and provides a general overview of activities in Norwegian public administration, again focusing on the perceived challenges of practitioners when attempting to embed datafication into practice, as well as issues drawn up in national policy and strategy documents. We focus especially on unseen issues in the discussion of the paper. Paper 3 is closely associated with the second paper, as it has its origins in one of the omitted issues identified in the analysis of Paper 2. Here, the policy production process and the resulting discourse in policy and among practitioners is in focus. We investigate how citizens and civil society are included and problematized in this paradigm shift in public administration and find a top-down and paternalistic approach to datafication in the Norwegian public sector. Finally, Paper 4 further investigates the sociotechnical imaginaries of public administration datafication through an in-depth analysis of policy papers. Here, the justifications of datafication are central to deepening our understanding of how the sociotechnical imaginary of data as a resource is framed and legitimized, but also how it changes over time.

Inspired by the turn toward collaborative ethnography (Lassiter 2005; Lassiter and Eric 2005), my research findings have been presented and discussed with the people and organizations working with and through data. At the intersection of all four papers, I develop the idea of what I call an "actionable critique of public administration datafication," which I present in the final chapter of this dissertation. The aim of the situated accounts of datafication is to emphasize the social dimension of data-driven technology, its relation to dominant agendas, and the potential for resistance (Dencik 2019). In doing so, I argue against a deterministic understanding of the data-driven society. This is of value not only to the scholarly debate on datafication but also to the practicing field of public administration and society as a whole. Although the contribution of this dissertation to the practicing field is not formalized in a published paper, insights from this research project have been continually fed back into the practicing field throughout the course of the last four years. I do regard the dissemination work done outside of the traditional channels of academic publication as highly important to the overall aim of the research project. Therefore, a list of activities can be found in Appendix B.

1.6 Structure of the dissertation

This thesis is article-based and consists of four research papers (three co-authored, one single-authored). This introduction to public administration datafication connects the papers and provides an overall account and discussion of the research project. Chapter 2 shortly reviews different terminology used to describe current developments and introduces the core concepts of "data-driven public administration" and "public administration datafication," as I regard the choice and definition of the concepts themselves as an important contribution to critical data and algorithm studies. This is followed by an introduction to the overall theoretical framework in Chapter 3. Chapter 4 outlines the research design of the dissertation and offers the reader in-depth reflections on the methodological choices made. Chapter 5 briefly introduces and summarizes the four papers, highlighting their key findings. The contributions of the overall research project are then presented and discussed in Chapter 6, where the key concepts of silences and challenges are especially in focus. This synopsis then culminates in some reflections on how critical data and algorithm studies might work toward an actionable critique of public administration datafication. The four papers are fully attached and followed by the appendix. Together, this narrative and the four papers presented aim to produce a compelling and thought-provoking deep dive into the inner workings of public administration datafication in Norway.

2 Clarification of concepts

This section explains the choice of concepts and their implications for the research project more generally. It serves as an extension of the previous presented literature and an introduction to the theoretical framework presented in Chapter 3. As Crawford (2021:7) points out, the multiple ways in which we define technology "is doing work, setting a frame for how it will be understood, measured, valued and governed." The literature reviewed in the previous chapter makes use of a variety of terms to describe what is researched, ranging from "digital transformation" over "algorithmic governance" to "automated decision-making." I have used the two key concepts of "data-driven public administration" and "datafication of public administration" across all four papers to describe what is researched and to set the scope for the project. I aim to use these concepts as boundary objects in the project, an "object which [is] both plastic enough to adapt to local needs and the constraints of several parties employing them, yet robust enough to maintain a common identity across sites" (Star and Griesemer 1989:293). Interdisciplinary research aiming to interact with a practicing field always operates in multiple social worlds and is, therefore, key to introducing concepts that manage to bridge these social worlds and facilitate communication between them. Indeed, the overwhelming and often unclear use of concepts in critical data and algorithm studies presents a key challenge to the scholarly debate and its dissemination to the practicing field (Moats and Seaver 2019). A short review of terms used, therefore, deserves its own section in this dissertation to provide the reader with a brief overview of overlapping and intersecting terminology in a heterogenic and emerging literature on algorithms, data, and AI technology in the social sciences.

2.1 Algorithm, big data, artificial intelligence

First, we might take a short look at the technical terminology present in both the academic field and the practicing field. The most common definition of "big data" refers to the three Vs: volume, velocity, and variety (Diebold 2012). Kitchin (2014b:68) then adds to this definition by adding the characteristics of exhaustive in scope, fine-grained in resolution, relational in nature, and flexible. The rapid growth of data in the last decade is often associated with the simultaneous development of technologies, infrastructures, and processes that make it possible to embed big data in practice. This includes developments within computing, networking, and storage. However, as scholars such as boyd and Crawford (2012) and Dencik (2020) point out, focusing only on the technical characteristics of big data often reduces our research focus to big data as a technical object driven by technical opportunities.

Within computer science, algorithms are defined as "any well-defined computational procedure that takes some value, or set of values, as input and produces some value, or set of values, as output [...] an algorithm is thus a sequence of computational steps that transform the input into the output" (Cormen et al. 2009:5). Computer scientists regard algorithms as tools to solve well-

specified problems. As Gillespie (2014a) points out, the term "algorithm" seems to be used by a variety of actors in different ways. For technical communities, algorithms are a simple technical solution for a technical problem of organizing data to achieve a specific outcome. For social scientists and the broader public, the term is often used to describe opaque artifacts that order our social lives. The terms "artificial intelligence" and "machine learning", used to describe complex forms of algorithms, might further complicate matters (Elish and boyd 2018). There is no widely accepted definition of artificial intelligence in the technical community (Wang 2019). This is then challenged by social scientists such as Crawford (2021:8), who argue that AI is neither artificial nor intelligence but embodies a material made from natural resources, human labor, infrastructures, and classifications, again pointing toward the wider political, material, and social structures that technology depends upon. Researchers such as Seaver (2017:1) therefore argue that social scientists should not align their use of the terms with the common definition in computer science, but "approach algorithms as 'multiples'—unstable objects that are enacted through the varied practices that people use to engage with them." The focus on practices will present a reoccurring theme in this dissertation.

As elaborated upon in Chapter 3 of this dissertation, I aim to decenter technology in the analysis of my papers. There has been an overwhelming focus on data as a technical artifact, decoupled from the social, in research (Dencik 2019). Simply stating that one researches algorithms, AI or big data runs the risk of giving a simplified and technocentric account of the phenomenon, fetishizing algorithms and code and disregarding the social and political aspects of the phenomenon (Elish and boyd 2018; Ziewitz 2016) . Because I realized during the initial field work in the Labor and Welfare Administration that the developments observed were about more than simply producing technology, I was in need of a different concept to describe what was researched and provide the scope for this dissertation.

2.2 eGovernment, digital transformation, and digital era governance

As indicated in the introduction, there is surprisingly little interest in big data, algorithms, and the associated technologies in public administration, with several of the top-ranked journals in the field only recently publishing work on big data and AI technologies (see, for example, Giest and Klievink 2022). These studies often adapt the technical fields' terminology and a technocentric approach to the phenomenon, studying these technologies in isolation. Research on technology in public administration seems to have been of limited influence on the academic field, as well as the practicing field, often kept separate from other core debates in public administration. As Meijer (2007:238) puts it, "Technology is complicated, not 'sexy' and not politicized and, therefore, not considered to be a core issue for administrators and politicians." However, this does not mean that digital technology is absent from the literature. In the field of public administration, itself an interdisciplinary subfield within the social sciences, the most common terms to describe socio-

technical developments within the public sector include “eGovernment,” “digital transformation,” and “digital era governance.” These concepts are interrelated, operating at the intersection of technology and public administration and often highlighting the potential of information and communication technology (ICT) to transform both services and public sector management (Mergel, Edelmann, and Haug 2019).

One of the first and most prominent concepts introduced in public administration studies to describe technological developments in public administration is eGovernment (Moon 2002). “Electronic government (eGovernment) is the delivery of services to citizens via the Internet” (Esteves and Joseph 2008:118). Research on eGovernment often focuses on citizen-government relations, the dissemination of technology and innovation, and the structural transformations of governance in the digital age, enabled by the internet (Heeks 2006; Meijer, Bolívar, and Gil-Garcia 2018). eGovernment is here praised as making public access faster and providing better information to citizens through internet websites. In its initial form, eGovernment research was mostly concerned with outward-facing services for citizens, but it has also been used more broadly to describe the introduction of ICTs into public administration (Meijer and Bekkers 2015).

A more recently introduced concept is the term “digital transformation,” which, according to Mergel et al. (2019), has its origins in the business sector but lacks a clear definition in public administration. This concept is often used to describe a more comprehensive organizational approach to digital technology and, therefore, goes beyond the mere digitalization of processes and services, encompassing both internal and external change. The transformative and unique capabilities of digital technology are also central to Dunleavy et al.’s (2006:478) notion of digital era governance:

> By digital-era governance we signify a whole complex of changes, which have IT and information-handling changes at their center, but which spread much more widely and take place in many more dimensions simultaneously than was the case with previous IT influences.

Digital era governance often focuses on organizational culture changes in public administration as a reaction to the shortcomings of new public management (NPM). Here, the reintegration of services, simplifying the relationship between public administrations and their clients, and an increase of productivity through IT and the internet, are central ideas. Margetts and Dunleavy (2013) also argue that a second wave of digital era governance is currently sweeping over public administration, enabled by social media and risk-based approaches to public governance.

Interestingly, research using these concepts often distances itself from technological determinist accounts of ICT, highlighting the multiple ways these developments are integrated in public administration and how they interact with, for example, organizational and budgetary factors, as well as focusing on people and culture (Alcaide–Muñoz et al. 2017; Shouran, Priyambodo, and Rokhman 2019; Dunleavy and Margretts 2013) . Nevertheless, Meijer and Bekkers (2015) stressed that research on ICTs in public administration is still dominated by explaining eGovernment rather than trying to understand eGovernment. The literature is thus dominated by a focus on

finding the key variables to determine the success and maturity of eGovernment (for example, using Rogers' (1995) diffusion of innovation perspective), rather than focusing on the socio-technical processes in specific contexts. In addition, this body of research is often overly optimistic in its view on ICTs in public administration (Heeks and Bailur 2007). Although researchers such as Dunleavy et al. (2006) warn about the possible pushback of digital era governance due to the possibility of linking data and enhanced tracking of citizens through their phones, highlighting the surveillance potential of digital technologies, these critical accounts of digital technology in public administration often remain a side note. The focus is on the enabling technology and optimistic outlooks of enhancing citizen trust and improving government performance through digital technology, rather than critically assessing its impact on society (Dobrolyubova 2021).

The literature associated with eGovernment, digital transformation, and digital era governance often includes all digital technology in its accounts of recent developments in public administration. Although these concepts help us understand the impact of ICTs within the wider landscape and history of public administration, I argue that they do not grasp the growing reliance of public administration on technologies such as algorithms and big data and the introduction of a new knowledge-making paradigm. These concepts seemed too broad and imprecise for the purpose of this dissertation. They focus on ICTs and technology more broadly and often disregard data as a central element of operation. Although some researchers have started to include the notion of data-drivenness in their accounts of the digital transformation of public administration, this remains an underexplored phenomenon (Dobrolyubova 2021; Fredriksson et al. 2017). Indeed, Heeks and Bailur (2007) argue that research on eGovernment often purposely disregards its links to information management in public administration.

As argued in Paper 2, I regard it as important to make distinctions between the digitalization of public administration, where online forms and internet communication are central, and recent developments in the public sector, which increasingly rely on data recirculation and the idea of data as a resource for and by public administration. In other words, datafication (which is introduced in section 2.4) builds on eGovernment and digital era governance efforts in the public sector, I do however want to argue that it presents a new paradigm for public administration; from digital to data-driven. It was, therefore, of importance to clearly distinguish my research object from the more commonly used terminology in public administration research, while simultaneously encouraging a dialogue between the field of critical data and algorithm studies and public administration. There is clearly a common link between my research and public administration literature in topics such as social equity, citizen participation, discretionary decision-making practices, and democratic governance. I regard publishing Paper 2 in a public administration journal as being of great importance to engage in this dialogue.

2.3 Algorithmic governance, automated decision-making, and new public analytics

Moving on to concepts more closely associated with critical data and algorithm studies, the term "algorithmic governance" was introduced by Müller-Birn, Dobusch, and Herbsleb (2013) to describe a coordination mechanism in opposition to "social governance." It is used to highlight the way digital technologies produce a specific "form of social ordering that relies on coordination between actors, [that] is based on rules and incorporates particularly complex computer-based epistemic procedures" (Katzenbach and Ulbricht 2019:2). This concept's origins can be traced to STS and its focus on the reorganization of social and sociotechnical practices (see, for example, Bijker and Law 1994). Katzenbach and Ulbrich (2019) argue here that "governance" implies a multitude of social ordering with regard to actors, mechanisms, structures, degrees of institutionalization, and distribution of authority and, therefore, embraces social ordering that is decentralized and not state-ordered. Studies applying the term "algorithmic governance" often focus on how algorithms contribute to shifting and re-organizing social interactions and structures and their ordering effect in specific contexts (see, for example Coletta and Kitchin 2017; Larsson 2018; Müller-Birn et al. 2013; Smith 2020). At the same time, these studies are not mere technological determinist accounts of the power of algorithms, but they always highlight the social, cultural, and political forces that shape algorithmic systems. A particular line of interest within studies on algorithmic governance is centered around the role of platforms and social media, where the way social media platforms order and structure information is central and the associated risks are explored (Gorwa, Binns, and Katzenbach 2020; Just and Latzer 2017; Musiani 2013; Ziewitz 2016). Studies on algorithmic governance often center around the ideas of agency, autonomy, and the risk of opacity in algorithmic systems. There are two distinct branches within this body of literature, addressing governance by algorithms and the governance of algorithms. While the former confirms Katzenbach and Ulbrich's abovementioned account, the latter is mostly concerned with the regulation or auditing of algorithms in society through legal or ethical guidelines (Danaher et al. 2017).

According to Katzenbach and Ulbrich (2019), the benefit of using theconcept of algorithmic governance is that it brings together a diverse set of phenomena, discourses, and research fields, managing to identify key controversies and challenges in a digital society. Danaher et al. (2017) develop a comprehensive research agenda on algorithmic governance that has been of great inspiration to this research project. Indeed, one of the papers presented in this dissertation has been part of a special issue on algorithmic governance in context. Nevertheless, studies on algorithmic governance are often outward-focused, meaning that studies within this field often focus on the consequences and impact of algorithmic systems rather than their construction and framing, they focus on how algorithmic ordering governs society (Issar and Aneesh 2022). In addition, the term itself implies that this phenomenon centers on algorithms rather than data, which I wanted to avoid in this dissertation. Furthermore, I wanted to avoid associating my

research with the governance of algorithms field, which is heavily dominated by ethics and law scholars (D'Agostino and Durante 2018; Saurwein, Just, and Latzer 2015). Although this body of literature is closely associated with critical data and algorithm studies, it often has a narrow focus on law and ethical guidelines to "solve datafication" and mitigate harm.

Algorithmic governance is often closely associated with automated decision-making. Automated decision-making directs our attention toward the multiple ways in which data and technology are used to automate knowledge production. This term highlights the delegation of decisions to algorithmic systems. I emphasize the argument of Andreassen, Kaun, and Nikunen (2021:213) that "the term 'artificial intelligence' should mostly be replaced with that of 'automated-decision-making,'" as AI often remains a future imaginary. However, researchers such as Kuziemski and Misuraca (2020) seem to use these terms interchangeable. This term draws attention to algorithm-based automation as a key emerging feature of society and literature associated with automated-decision making does often discuss issues of automated vs. human decision making (see for example Ranerup and Henriksen 2022; Wagner 2019). My work is in addition inspired and complements work done in the 'Automated decision-making: Nordic Perspectives' network that encourages interdisciplinary and empirical work on automated decision making in the Nordics (University of Copenhagen 2022).

The focus on knowledge production through algorithms is also apparent in Karen Yeung's (2020) notion of new public analytics, a term she uses to describe how predictive analytics have become a key element of public administration. Yeung argues that new public analytics presents an administrative reform to the same extent that new public management (NPM) did. While NPM introduced new forms of management to public administration, new public analytics introduces new forms of knowledge and automated decision-making into its inner workings. Yeung characterizes this as a distinct break with the past, where algorithmic ordering and the professions of data science come to dominate administration. This once again reinforces the argument that datafication brings with it a distinct change and must therefore be seen as something other than eGovernment or Digital Transformation.

Automated decision-making and predictive analytics are clearly a part of the imaginary of the data-driven public sector. I do however want to argue that these concepts too narrowly focus on the automation of decision-making processes or introduction of prediction and are not able to sufficiently grasp what is currently empirically observable in the public sector in Norway. This is not just about automation or just about prediction, but a powerful coming together of a variety of ideas aiming to fundamentally change how the public sector operates. I wanted to draw attention to data use and recirculation in public administration and did therefore decide to make the datafication of public administration the main object of inquiry of my dissertation. It is the intersection of a variety of processes and practices that makes up public administration datafication.

2.4 Core concepts: Datafication and data-driven public administration

This research project makes use of the term "datafication of public administration" to encompass a variety of processes, technologies, and practices in the public sector. The choice of terminology is closely associated with the empirical work done in the dissertation. In the project proposal of this dissertation, I made use of the terms "AI" and "algorithmic governance" to describe what was researched. However, following my micro-level study of the Labor and Welfare Administration's AI lab in 2018 (Reutter, 2018), I realized that my research focus was too narrow. I was searching for AI and algorithms but seemed to be missing the bigger picture. Something was amiss—this was not just about automation and AI. The research project was, therefore, in need of a broader scope and a more precise terminology to be able to grasp what is observed.

The current development in the public sector was as much about data and the recirculation of data as was emphasized in the introductory quote: "Facebook earns big money with insights based on their users' data. What can Norway learn from Zuckerberg and co.?" The empirical findings of all the papers presented in this dissertation do indicate a strong focus on changing data practices, in addition to the introduction of algorithmic systems and automated decision-making into public administration. It is these changing data practices that are the focus of the presented work, and the chosen terminology reflects this focus and a system-level approach to the research object. Using the term "datafication" also allows me to directly position my work in relation to researchers who have been of great influence and inspiration to this dissertation by sharing this common terminology (Dencik 2019; Dencik and Kaun 2020; Hintz, Dencik, and Wahl-Jorgensen 2018; Redden 2018; Redden et al. 2020). At the same time, my understanding differs in some ways from these scholars' notions of the datafication of the public sector, which I will elaborate on here.Indeed, the work done on broaden the scholarly understanding of public administration datafication can itself be seen as a contribution of my dissertation.

The concept of datafication was introduced by Mayer-Schönberg and Cukier (2013:78), who stated that "to datafy a phenomenon is to put it in quantified form so that it can be tabulated and analyzed." The authors emphasized that datafication is about reducing information to elements that can be processed by computers and algorithms, thus focusing on processes of "dematerialization." Datafication builds on digitalization but needs to be differentiated. This definition is further extended by Mejias and Couldy (2019:2): "datafication combines two processes: the transformation of human life into data through processes of quantification, and the generation of different kinds of value from data." While Mayer-Schönberger and Cukier (2013) and Mejias and Couldry (2019) mostly focus on a general datafication of social activity into analyzable form, I want to differentiate this understanding of datafication in my work to situate datafication in the context of public administration. Like boyd and Crawford (2012), I regard datafication as a phenomenon that rests on the interplay of technology, data, and the idea of data as the enabler of a good society. The datafication of public administration is, therefore, about the

introduction of new technology, data, and the idea of data as a resource for improving public governance. The datafication of public administration indicates a process, a continual change of public administration practices, rather than a single technology or a one-time event. This process goes beyond the individual and confined applications of the delegation of responsibility to algorithmic systems or automated decision-making to encompass a wider scope of practices. It aids the dissertation to direct its analytical attention toward a paradigm shift across the public sector, where data recirculation and extraction have become a key feature.

While I use public administration datafication in the scholarly dissemination of my research, the practicing field describes its own practice as "working towards a datadriven public administration" (Cavanillas, Curry, and Wahlster 2016). According to the OECD, "A Data-Driven Public Sector (DDPS) transforms the design, delivery and monitoring of public policies and services through the management, sharing and use of data" (van Ooijen, Ubaldi, and Welby 2019:6). The term "data-driven" denotes that something is driven by data, again indicating the central idea of data as an enabler of public administration, something that directs the public sector. The OECD definition of data-driven public administration is quite vague, and the findings of Papers 1–3 indicate that there is a lack of common understanding of data-driven public administration among practitioners and policy makers. Therefore, I saw a need to develop my own definition of data-driven public administration for the purpose of this research project:

> Data-driven administration consists of two interwoven processes: the use of more and/or different kinds of data and the recirculation of data through ever more complex methods, (such as machine learning or automatic decision-making).

Like Mejias and Couldry's (2014) definition, this definition describes datafication as a two-fold process. Data-driven public administration is a socio-technical phenomena that is materialized in a variety of efforts today, ranging from calls for open government data to the establishment of analysis platforms in public administration and the development and implementation of automated decision-making and machine learning in case work. This definition encompasses both data and technology at the same time, drawing attention toward the procedural nature of this development by addressing the underlying idea of the datafication of public administration. Note that data is not directly linked to "big" and, therefore, encompasses both small and big data (Kitchin 2014). I regard all technology or systems that makes use of any kind of data and which focus on the recirculation of data as data-driven technology. This includes simple, rule-based automatic decision-making, as well as more complex technologies, such as machine learning and AI. Again, datafication is not about a single technology or knowledge paradigm. It focuses on data as a resource as its core idea. This also directly connects to the work of Dencik et al. (2019) on the scoring society, where prediction, classification, and scoring of citizen activities is central.

Many researchers argue that the novelty of datafication lies in "quantifying elements of life that until now were not quantified to this extent" (Mejias and Couldry 2019:3). While I regard the act of quantifying as inherent to datafication and a key characteristic of datafication, this

understanding is in need of an alteration in the Norwegian public administration data context. Data and the modern state are inseparably woven together as the availability of statistical information is of key importance to state governance (Desrosières 1998). Quantification has always been part of public administration (Porter 1996). We cannot decouple previous data practices from the current datafication of public administration as this would obscure the underlying power dynamics of datafication (Cieslik and Margócsy 2022). However, the speed, scale, and ubiquity of data practices are changing (Danaher et al. 2017). The datafication of public administration is not only about an increase in the quantification of every aspect of social life but also about the recirculation of data that has previously had only very limited use, such as register data. Register and other kind of administrative data in Norway has mainly been used for statistical purposes and research, but as I argue in a 2019 book-chapter, is now imagined to be actively used in the day-to-day operations of public administration (Reutter and Spilker 2019). Therefore, the first part of the datafication definition introduced here is about the "use of more and different data." I deliberately chose the term "use" and not "collection" or "quantification" as many of the datafication projects observed do not rely on the collection of new data but are about actively engaging with data already collected and stored within public administrations. Public administration has already quantified many aspects of our social lives. The emphasis, then, is on "more" and "different," indicating that datafication changes the way data is used, what data is used, and to what extent. Furthermore, recirculation is about an activization of data and the act of feeding data into practice. While researchers such as Mejias and Couldry (2019:7) argue that "The analytical value of the term 'datafication' lies in its ability to name the process and the frameworks by which a new form of extractivism is unfolding in our times, via the appropriation of data about our lives," I did consciously choose to not use extractivism in my definition. This is clearly part of the datafication paradigm of public administration, as shown in Paper 4 and argued by Dencik (2022), however dafication is not confined to extractivism

"Can we save lives, get people into jobs and education, and create a better society? Yes of course!" The second part of the introductory quote indicates a growing trust in data as an enabler of the good society; this is what van Dijck (2014) calls "dataism." "[The] datafication paradigm therefore relies on particular epistemological and ontological assumptions, underpinned by its own specific set of values and logics—and politics" (Dencik 2019:244). Datafication is fueled by the idea of data as a resource for knowledge production that is superior to other forms of knowledge production (McQuillan 2018). Couldry argues that "what counts as social knowledge, and who/what counts as an input to social knowledge, is changing" (2020:1139). This may alter our understanding of society as a whole. As van Dijck (2014) pointed out, the prioritization of data over other inputs to social knowledge is beginning to change how we know the world and how we approach the social. This reflects Dencik et al.'s (2019) idea of citizen scoring enabled by datafication that fosters new ways of categorization, assessment, and prediction at both the individual and population levels.

In my papers, I have described datafication as an administrative reform. This was a strategic classification to direct the attention of public administration scholars toward datafication, at the same time as position data-driven processes in relation to previous public sector reforms, such as eGovernment (Kraemer and King 2006). Administrative reform is a well-known concept within public administration studies and can again act as a boundary object. As argued , public administration scholars have often focused on big data and AI in isolation rather than discussing the system-level and societal issues associated with datafication (Agarwal 2018; Klievink et al. 2017; Maciejewski 2017). Administrative reform is defined by Montgomery (1967:1) as "a political process designed to adjust the relationships between a bureaucracy and other elements in a society, or within the bureaucracy itself." This definition stresses the procedural nature of datafication, as well as its close connection to politics and power. As Dencik and Kaun (2020) argue, the datafication of the welfare state needs to be regarded as a political development rather than a simple matter of efficiency or quantification and regarding data-driven public administration as an administrative reform then draws our attention towards the relations between administration, politics, and society. At the same time, I have here attempted to study an administrative reform not simply through the lenses of public administration scholarship (Caiden 1999; Christensen, Lie, and Lægreid 2008; Cordella and Bonina 2012), but focused on socio-technical approaches and practice theory to understanding datafication, as introduced in the next chapter.

In summary, the two key concepts of data-driven public administration and public administration datafication provide the scope for this research project. By advocating for this terminology, I have already positioned my research project within a specific ontological and epistemological understanding of what was researched and how. Public administration datafication is seen as a continuously constructed project that relies on the interplay of data, technology, practices, and an overall belief in the superiority of knowledge produced through data. It is a sociotechnical phenomenon by definition; one that promotes the idea of data as a resource for improving public administration and society more general. How then can we research public administration datafication? This leads us to the theoretical framework of the dissertation.

3 Theoretical Framework

> Regardless of the lens, what is required is deep, careful and critical reflection and putting theory to work through empirical case studies (Kitchin 2014:186)

This dissertation takes Kitchin's suggestion seriously, putting theory to work through empirical case studies in a creative and explorative way. The theoretical framework presented in this section serves as a blueprint for the dissertation and provides the structure for how the overall research question has been philosophically, epistemologically, methodologically, and analytically approached. Using an abductive approach to public administration datafication required me to reflexively engage with existing theoretical reasoning on technology in society, while trying to make sense of the empirical observations in the field. Research on datafication is interdisciplinary in nature and therefore allows us to engage creatively with a variety of scholarly thinking and entry points (Lomborg, Dencik, and Moe 2020).

In introducing the theoretical framework, I argue that there are four central puzzles social scientists are facing when attempting to produce critical accounts of public administration datafication. These puzzles have their origin both in the theoretical and scholarly debates in critical data and algorithm studies (Couldry 2012; Dencik 2019, 2020; Kitchin and Lauriault 2014; Seaver 2019) and in the tensions that arose in trying to make sense of what I was observing in the field and matching these observations with existing literature on STS, algorithmic systems, big data, and automated decision-making.

Puzzle 1: How to avoid determinism in the analysis of data-driven public administration.

Puzzle 2: How to navigate between hype/discourse and real world/materiality in studying public administration datafication.

Puzzle 3: How to navigate between critical accounts of specific applications of data-driven technology in the public sector and a general critique of the datafication of public administration/society.

Puzzle 4: How to balance between proactive and reactive critiques of datafication.

I will use the next pages to elaborate on these puzzles. This reflects my thought processes in working on this dissertation and provides valuable insights into how and why I have studied public administration datafication the way I have. In the search for a way to approach these puzzles, I have developed the idea of an actionable critique of public administration datafication, which I regard as a key contribution to critical data and algorithm studies derived from the intersection of the four papers and which I will elaborate upon in the final section of the dissertation. Although each paper makes use of a different analytical framework to address the problems presented, they share a common foundation from which knowledge is constructed and, therefore, share a system of concepts, assumptions, and beliefs that guided the overall research project (Miles and Huberman 1994). To bridge all four puzzles, I combined practice, theory, and STS, in addition to

revisiting critical theory. I will also introduce two key concepts: the data assemblage and sociotechnical imaginaries that have been important both as methodological lenses and as analytical tools in this research project.

3.1 Studying data-driven technology in the social sciences: From critical theory to STS and media practice

In this first part of the theory chapter, I outline what Timmermans and Tavory (2022:9) call "landmarks of expectation against which new theoretical paths may be tracked." In their account of abductive analysis, the authors stress the importance of a broad understanding of social theory when making use of abductive analysis to facilitate moments of surprise. These moments of surprise can then be used to work toward theorizing. We therefore start broad in laying out the theoretical framework of the dissertation. As argued above, critical data and algorithm studies is a loosely connected and emerging research field. I want to argue that we can trace back the origins of this field to two distinct lines of thoughts: critical theory and STS. Many of the key assumptions in the field have their origins in these scholarly traditions, and they are therefore in need of a quick review. Rather than taking a deep dive into critical theory and STS, I will just scratch the surface to quickly move on to practice theory, as introduced by Couldry (2004), to begin positioning my own work in relation to the previously introduced puzzles.

Critical theory plays a central role in the social sciences when discussing basic sociological concepts such as rationalization, capitalism, power, and ideology in relation to technology. It offers a valuable contribution to the critique of positivism (Agger, 1991). One of the key sources of critical theory is the Frankfurt School, a collection of neo-Marxist scholars established in the 1920s. It is an interesting point of departure to understand the foundation of critical accounts of technology in society in the social sciences. The Frankfurt School adopted Weber's view of modernity as defined by differentiation and the fragmentation of the world into spheres of value (Kirkpatrick 2017). The technological sphere is autonomous in the hands of experts and represents a dominating force in this view. Society is increasingly rationalized, which leads to technocratic thinking and a loss of reason. Reason assesses the means to ends in terms of ultimate human values such as justice, happiness, and freedom, instead of the technocratic value of efficiency (Beira and Feenberg 2018). Similar thoughts can for example be found in Issar and Aneesh (2022) account of algorithmic governance, where the authors argue that this development decreases spaces of negotiation. The continued expansion of rationalization in the culture, technology, and knowledge industries leads critical theorists to a pessimistic view of the future, with little room for configurations. Following this line of thought would then lead us to expect datafication to enter and change public administration uninterrupted, expanding the administration's abilities of power and control and the increasing rationalization of the sector.

Technology was a central object of discussion for early Frankfurt School scholars such as Heidegger, Adorno, and Marcuse. In his 1941 article "Some Social Implications of Modern

Technology," Herbert Marcuse argued, for example, that technology in the contemporary era constitutes an entire "mode of organizing and perpetuating (or changing) social relationships, a manifestation of prevalent thought and behavior patterns, an instrument for control and domination" (Marcuse, 1941:414). This is elaborated in his book *One-Dimensional Man* (Marcuse, 1964), where he argues that modern technology (in his case the television) advances repression, limiting the ability to think critically about society and technology. I find this a highly interesting book, as many of the arguments resemble recent critiques of datafication. Heidegger (1977) then regarded technology as a dominating and controlling way of thinking and engaging with the world. Habermas, on the other hand, represents a later version of the Frankfurt School. Although Habermas pointed out that "social interests still determine the direction, function, and pace of technical progress" (Habermas 1970:105), he also regarded scientific-technical rationality as non-social, neutral, and formal. Technology by definition excludes the social. Science and technology do not respond to social interests but only to the objective worlds that they represent in terms of the possibilities of understanding and control. Technology is part of the system world and is not accessible to meaning-oriented interpretation and critique. If technology is part of the system world and not accessible to meaning-oriented interpretation, how is a critique of technology possible for social scientists?

Control, power, ideology, and rationalization are central objects of investigation in critical data and algorithm studies (see, for example, Beer 2017; Eubanks 2018; McQuillan 2016; O'Neil 2016; Van Dijck 2014). As with the Frankfurt School, a critique of positivism is at the heart of this line of thought (boyd and Crawford 2012; Dalton and Thatcher 2014). However, while critical theory reminds us of critically interrogating and questioning the taken-for-granted ideas and values of technology, the Frankfurt School often presents both a technological deterministic view and substantivist views on the relationship between technology and society (Kirkpatrick 2017), which I wanted to avoid in this dissertation. In addition, this line of thought has been heavily criticized for producing too abstract and pessimistic accounts of technology and society (Feenberg 1991). Following my overall ambition to produce critical accounts of datafication, that can be fed back into practice, therefore required me to look elsewhere. A reaction toward the deterministic and often very theoretically oriented work of the Frankfurt School on technology can be found in STS.

STS is an umbrella term for a variety of theories and methodological approaches to studying technology and knowledge in society. Giving a full account of this line of research would extend the frame of this dissertation. Critical data and algorithm studies have clear ties to STS and its interest in complex technical systems (Hughes 1993; Winner 1993) and the critique of rationalization, classification, formalization, and quantification produced by scholars such as Bowker and Star (1999) and Porter (1996). In its early days, STS was characterized by studies of technology and knowledge in the making (Callon 1984; Latour and Woolgar 1979; Latour 1987; Woolgar 1990), while this has been supplemented with a user focus in more recent work (Silverstone and Haddon 1996). STS mirrors many of the studies associated with critical data and

algorithm research, often focusing on demonstrating bias (race, gender, class) or danger (nuclear waste, recombination DNA) in scientific work (Star 1988).

In particular early STS has provided some of the cornerstones for this dissertation, promoting the idea of the social construction of technology and data and the importance of both historical accounts of emerging technology and ethnographic studies of technology production (Bijker and Law 1994; Hughes 1993; Latour 1987; Pinch and Bijker 1984). I was highly inspired by this work when embarking on my first explorative field work in 2018. The social constructivist view provides the very core of the project, as opposed to the deterministic and substantivist understanding presented by critical scholars associated with the Frankfurt School. As Seaver (2019:413) pointed out, "The point of declaring something a construction is to argue that it might be constructed differently." However, STS-informed scholarship has also been heavily criticized for producing far too descriptive accounts of technology in society and often disregarding aspects of power and ideology (Feenberg 1991; Kirkpatrick 2017). Therefore, I went in search of an approach that helps me balance STS and critical theory. A variety of scholars have attempted this, in my theory-matching process, I, however, repeatedly turned back to practice theory as introduced by Couldry (2004).

Couldry (2020:1141) argues that critical work within the social sciences addressing datafication relies too much on the theoretical resources shaped by the legacy of STS and its flat ontology. This fails to answer the questions of

> How is the overall order of social life being reconfigured to promote particular corporate and governmental interests on the basis of new and radical forms of reduction—the reduction of human life to configurations from which profit through data can be maximally extracted.

He therefore suggests that critical scholars of datafication need to redirect their attention toward asking questions about how particular assemblages of datafication emerge and stabilize and toward the larger social and economic forces that shape social order. This might require us to revisit social theory beyond STS. While Couldry draws upon researchers such as Norbert Elias, Luc Boltanski, and Judith Butler in his paper, I directed my theoretical ambitions toward Couldry's own work on practice theory (Couldry 2004, 2012). Practice theory is a theoretical turn in the social sciences that attempts to overcome theoretical divisions between structure and agency and individual versus society. It consists of a vast and heterogenous literature, stretching from Schatzki to Bourdieu and Foucault. Practice theory sees practice as the essential atom of the social world and encourages researchers to investigate the mechanisms, interactions, and interconnections that occur between and among practices (Bakardjieva 2020:2933) .

Drawing on Ann Swidler (2001), Couldry (2004) integrates practice theory into media sociology. According to him, this helps anchor media science more closely within the social sciences. As shown in the introduction, many scholars associated with critical data and algorithm studies have close ties to media studies (see, for example, Bucher 2012; Dencik 2020; Gillespie 2014b). While much of media science has been directed toward studying audiences or users and media effects,

practice theory allows us to broaden our perspective and study the whole range of practices related to media (Couldry, 2012). It decenters the media text itself and therefore avoids making claims about its immediate effects. Practice theory therefore seemed to be a valuable landmark of expectation to explore in my quest for a theoretical framework to accompany this dissertation. At its most basic level, practice theory invites us to ask, "what are people (individuals, groups, institutions) doing in relation to media [technology] across a whole range of situations and contexts" (Couldry, 2012, chapter 2, section 2).

I want to follow Bakardjieva's (2020) suggestion to use practice theory as a "sensitizing concept" that "gives the user a general sense of reference and guidance in approaching empirical instances" (Blumer 1954:6) rather than prescriptions of what to see. Practice directs our analytical attention toward the practices of actors and discourse (Couldry, 2012). It highlights how certain practices relate to other social practices. By emphasizing practice, we overcome the determinist ideas of technology steering society. It also helps us to translate hype into more concrete questions of practices (Couldry 2012) and therefore also helps us to approach puzzle 2. At the same time, this framework encourages an openness to the varied and complex organization of practices and how they are ordered without disregarding power dynamics altogether. In bridging practice theory with STS, I am therefore able to produce a situated account of datafication in the context of public administration.

3.2 Toward a situated practice approach

Couldry (2020:1146) regards the challenge of the social sciences today as an effort "to understand the form and dynamics of processes of datafication, and their consequences for wider social orders that characterize the contemporary social world." It is therefore important to ask through what institutional action and material resources new social orders emerge and how larger forces shape datafication and its resulting forms of power. Couldry's work on media practice inspired Dencik (2019, 2020) in her call to integrate practice theory in studying datafication, which has been central in this dissertation and provides the core of this theoretical framework. It also helps us position the papers in relation to several of the previously mentioned puzzles.

In their recently published literature review on datafication, Flensburg and Lomborg (2021) stress that this scholarship's interest can be grouped into two distinct groups: user understandings and user practices and infrastructure and technological processes. As Dencik (2019) pointed out, there has been an overwhelming focus on data as a technical artifact, decoupled from the social, in research on datafication. Focusing on technical issues and technical modes of bypassing harm, this body of literature mostly focuses on the phenomenon in general and issues such as the lack of transparency, bias, and discrimination, as argued earlier in this dissertation. Drawing on Christin (2017), Dencik (2019) argues that most of the discussion around datafication has focused on data and algorithms and their functions rather than on the practices, representations, and imaginaries of the people who construct and rely on data systems in their everyday lives. The

discussions in the field have often neglected the social and political dimension of data and technology and the processes that lie behind working systems. “A critical practice approach to data constitutes a powerful lens to overcome the prominent data centrism in studies of big data and algorithms, restoring the political dimension of datafication” (Stephansen and Treré 2019:16). Drawing on Couldry (2012), Dencik (2019) advocates a situated practice approach to datafication in which we focus on what people do in relation to data and technology in the contexts in which they act and, therefore, actively decentralize data and technology from our analysis. This invites researchers to “uncover key questions about the values and interests that pertain to data in different contexts” (Dencik 2019: 243).

Couldry and Powell (2014:1) argued that datafication should be investigated in a way in which the agency and reflexivity of individual actors is foregrounded, “as well as the variable ways in which power and participation are constructed and enacted.” Understanding datafication is about an open enquiry into what social actors and organizations do with and under a datafication regime. In this dissertation, public sector practitioners and policy makers in particular are in focus. The aim of the situated accounts of datafication is to emphasize both the social dimension of data and its relation to the dominant agenda and practices and potential for resistance. This invites researchers to focus on the underlying social mechanisms within that context in relation to agents working and living within such contexts, giving them the opportunity to reject datafication as an inevitable development and, instead, see it as a “continuously constructed project, shaped by multiple, converging and conflicting forces” (Dencik, 2019:246). In other words, it helps us to avoid determinism as problematized in Puzzle 1 when studying public administration datafication and refocus our scholarly attention toward how dominant agendas and practices might be altered. This also resonates Issar and Aneesh (2022) argument that what can bring together the diverse set of social science inquiries into data and algorithms is a focus on ‘negotiability’ and spaces of negotiability.

There is a distinct lack of research addressing the context dependency of the datafication paradigm as most researchers focus on the universal characteristics of datafication (Wijermars and Makhortykh 2022). How then can we navigate between critical accounts of specific applications of data-driven technology in the public sector and a general critique of the datafication of public administration in our research? Public administration consists of a distinct institutional action and material resources. Drawing on Couldry (2012), Dencik (2019) stresses that we need to consider how distinct contextual practices stand in both tension and alignment with the datafication paradigm. Situated accounts of data-driven public administration take the contextual practices of public administration seriously and demonstrate that datafication might be embedded and framed differently in public administration than, for example, on internet platforms. By paying attention to the perceived obstacles and challenges of practitioners working on and through data-driven systems, I attempt to highlight the multiple forces that shape public administration datafication, as well as making the often-invisible work involved in recirculating

data visible. In doing so, I focus on the contradictions, ambiguities, and tensions between datafication and other social practices, following Couldry (2004).

As Crawford (2021:9) argues, before understanding the societal implications of datafication, we need to "ask what is being optimized, and for whom and who gets to decide." We then come to the puzzle of reactive vs. proactive critiques of datafication. This puzzle is about the temporality of critique and the artificial scholarly distinction between the production and use of data-driven systems. Couldry (2004) and Dencik (2019) mostly focus on the users of technology or media and their potential to give datafication new meanings. Through my dissertation, I do, however, argue that the everyday experiences of people working on embedding and framing datafication in public administration are of great importance to advance our understanding of datafication as a practice. By paying close attention to the converging, and conflicting forces involved in the continuous construction and framing of datafication in public administration, we can critique current developments more proactively rather than solely targeting our critique toward existing and working systems. This focus on practitioners is also inspired by Seaver's (2017) call to dissolve the split between "insider" engineers that know about the functions and making of data-driven systems and "outsider" social scientists that know about its consequences. As empirically demonstrated in this dissertation, datafication does often remain a future vision. Imaginaries do have performativity, as argued later, but they do not always materialize without interruption. This is especially important as researchers approaching datafication empirically often need to navigate between hype and materiality.

The situated practice approach introduced in this section provides this dissertation with an overall shared set of concepts, assumptions, and beliefs that guided the research project. To operationalize this approach, I will furthermore introduce two key concepts, data assemblage and sociotechnical imaginaries, which have been of great use to the research project. Sociotechnical imaginaries allow us to understand the relationship between political power and technology in the promotion, production, and reception of data-driven public administration and therefore help us to understand what people imagine in relation to datafication. Data assemblages serve as both a methodological lens (further elaborated on in Chapter 4) and a practical analytical tool to approach datafication. It assists me in the operationalization of Dencik's work.

3.3 Data assemblages

Critical data and algorithm studies make the case that data-driven systems should be regarded as "always already constituted within wider data assemblages" (Iliadis and Russo 2016). This helps us capture the multitude of ways that already-composed data structures inflect and interact with society, its organization, and its functioning, and the resulting impact on individuals' daily lives. New data practices are always already embedded into existing practices, as pointed out by Couldry and Dencik. Kitchin (2014) therefore encourages one of the ways to enact critical data and algorithm studies is to focus attention on the socio-technical assemblages that produce,

circulate, and utilize data in diverse ways. This is a concrete way to think about datafication, which, as Couldry (2004) suggested, decenters the technology and instead pays attention to how data-driven systems come into being and are made up by a variety of elements. Kitchin and Lauriault (2014) argue that data systems are not interdependent of contexts, ideology, political interests, and economic forces. The data assemblage is bound together in a set of contingent, relational, contextual, discursive, and material relations and consists of more than data and algorithms themselves, once again reminding us of the importance of directing our analytical attention away from simply regarding code.

The framework of the data assemblage (see Table 1) as introduced by Kitchin and Lauriault provides this project with a grounded and situated analysis of the politics of data systems in the datafication of public administration. Again, this can be connected to moving beyond a simple critique of positivist knowledge production and toward more contextual and situated understandings of data in society. This also reflects Seaver's (2019) suggestion to approach data-driven technology as heterogeneous sociotechnical systems that are influenced by cultural meanings and social structures. This lens draws our attention toward flux, revisability, and negotiation, as demonstrated in all four papers. The technological, political, social, and economic apparatuses frame the assemblage's nature and work (Kitchin and Lauriault 2014). Therefore, data assemblages are the result of a diverse set of actors, institutions, technologies, practices, structures, governmentalities, and knowledges (Kitchin 2017). These apparatuses and their elements interact with and shape each other, and they frame what is possible, desirable, and expected of data and algorithms. Critical research on data-driven systems needs to examine the logics that guide the apparatuses of the system.

Kitchin and Lauriault (2014) were inspired by Foucault's idea of power/knowledge and the dispositive, a "thoroughly heterogenous ensemble" (Foucault 1989:194) in their work on data assemblages. According to Foucault, the dispositive of data infrastructures produces knowledge that fulfills strategic functions, stressing the importance of regarding data systems as entangled with power and as non-objective or neutral. Furthermore, they draw on Ian Hacking's (1986) work to argue that data assemblages are always part of a wider data landscape. "Within the public sector, for example, there are thousands of data systems, each surrounded by a wider assemblage, that interact and work in concert to produce state services and forms of state control at local, regional and national scales." (Kitchin and Lauriault 2014:10)

Table 1. The apparatuses and elements of the data assemblage (Kitchin 2014:25)

Apparatus	**Elements**
System of thought	Modes of thinking, philosophies, theories, models, ideologies, rationalities, etc.
Forms of knowledge	Research texts, manuals, magazines, websites, experience, word of mouth, chat forums, etc.
Finance	Business models, investment, venture capital, grants, philanthropy, profit, etc.
Political economy	Policy, tax regimes, incentive instruments, public and political opinion, etc.
Governmentalities and legalities	Data standards, file formats, system requirements, protocols, regulations, laws, licensing, intellectual property regimes, ethical considerations, etc.
Materialities and infrastructures	Paper/pens, computers, digital devices, sensors, scanners, databases, networks, servers, buildings, etc.
Practices	Techniques, ways of doing, learned behaviors, scientific conventions, etc.
Organizations and institutions	Archives, corporations, consultants, manufacturers, retailers, government agencies, universities, conferences, clubs and societies, committees and boards, communities of practice, etc.
Subjectivities and communities	Of data producers, experts, curators, managers, analysts, scientists, politicians, users, citizens, etc.
Places	Labs, offices, field sites, data centers, server farms, business parks, etc., and their agglomerations
Marketplace	For data, its derivatives (e.g., text, tables, graphs, maps), analysts, analytic software, interpretations, etc.

In this research project, I have made use of the idea of the data assemblage as both a methodological lens and an analytical tool. Placing Kitchin and Lauriault's idea of the data assemblage within practice theory, this dissertation asks how the apparatuses shape what people and organizations do in relation to datafication. In addition, I regard the process of assembling data assemblages as a distinct practice worth studying as it helps us understand how apparatuses and elements stand in both tension and alignment under the datafication paradigm. As Bakardjieva (2020:2941) highlights,

> […] new social phenomena […] arise out of the emergence of some "new combinations of doings and sayings, rules, teleologies, understandings, material arrangements, and relations between practices and arrangements" as Schatzki (2011:7) maintains, the goal of research is not simply to describe these new combinations, but to trace the factors and forces that make their emergence possible.

The apparatuses outlined by Kitchin and Lauriault (2014) can help us understand the factors and forces that make datafication of public administration possible and offer us a concrete analytical

tool to study the new combinations of the elements of the data assemblage. Furthermore, in my work, I have continuously focused on obstacles and challenges to the embedding of datafication in public administration, as especially apparent in Papers 1 and 2. reframing the apparatuses as constraints can help us consider how distinct contextual practices stand in both tension and alignment with the datafication paradigm.

What are reasons for subjecting a system to the logic of datafication in the first place (Kitchin 2017)? Here, the concept of sociotechnical imaginaries is central. I regard the system of the thought apparatus as a highly interesting object of inquiry, especially in this early phase of public administration datafication, and argue that we need to bridge the concept of the data assemblage and Jasanoff and Kim's (2009) idea of sociotechnical imaginaries to broaden our understanding of how the of the data-driven public sector is embedded and framed in Norway and the multiple practices that constitute datafication. As I pointed out in the introduction, most of the work done in critical data and algorithm studies is concerned with already-functioning and materialized systems. However, in many countries, what we deal with is future visions rather than actual working systems. I needed a way to address these future visions and understand their role in the datafication of public administration without simply denoting them as hype. This is where the concept of sociotechnical imaginaries was of great use.

3.4 Sociotechnical imaginaries

To contribute to the theory regarding the relationship between political power and science and technology, Jasanoff and Kim (2009) introduced the concept of the sociotechnical imaginary (STI). This concept serves to bridge studies on the production of science and technology with the promotion and reception of these by non-scientific and non-technological actors and institutions, such as public administrations and governments. According to Jasanoff and Kim (2009), STIs help us understand local science and technology variations in policy and practice. The authors define the concept as "collectively imagined forms of social life and social order reflected in the design and fulfillment of nation-specific scientific and/or technological projects" (Jasanoff and Kim 2015:120). As with Dencik (2019), this concept emphasizes the relevance of context and future vision. Jasanoff suggests that STIs tend to emerge in four phases, from their origin, to embedding, to resistance, to extension (Jasanoff 2015). In this dissertation, especially the first two phases have been of great importance. Thus, STIs offer a valuable conceptual lens for exploring how various actors produce future visions and expectations at the intersection of politics, discourse, and technology (Bareis and Katzenbach 2021). I regard the concept of STIs as highly valuable in addressing Puzzle 2, balancing between hype and materiality, and in my quest to produce proactive critiques of technology.

The concept demonstrates how visions and expectations of futures are always embedded in the social and institutional practices of science and technology and thereby influence technology development (Fujimura 2003; MacKenzie 1996). Again, this resembles Dencik's (2020) focus on

social, institutional, and organizational practice. The objective of STIs as a sensitizing concept is not to outline STIs in detail at a particular point in time but to observe how technological development processes are constantly shaped and in the making at multiple sites by multiple actors who share and stabilize the visions and expectations of the technologies, directing our attention to processes associated with assembling data in specific practices. In this dissertation, then, the practices of embedding, maintaining, circulating, and materializing STIs is in focus. Once sociotechnical imaginaries are embedded, employed, and adopted, they are seen as "natural, inevitable, or determined in advance" (Jasanoff 2006:277). They appear inevitable and researching them while they are embedded and maintained is therefore important.

STIs are future-oriented and often lack present materiality. Nevertheless, these visions provide the foundation on which policies are built and therefore influence and legitimate technology production. Drawing on Appadurai (1996) and Taylor (2004), Jasanoff and Kim (2009) argued that imagination is here regarded as "an organized field of social practice" and a key element in making social order rather than fantasy or illusion. It is important, then, that STIs have agency. This helps us to bridge hype and materiality as, even if they are not materialized in specific data assemblages, STIs have impact on current practices, such as funding allocation and the establishment of data teams across the public sector. As Jasanoff and Kim (2009:120) point out, they have "the power to influence technological design, channel public expenditures, and justify the inclusion or exclusion of citizens" and are therefore important objects of investigation, even though they are promises and visions rather than specific technologies. This concept is then important to help us navigate between hype and visions and material infrastructures.

A variety of researchers have begun to investigate the STIs of datafication (Bareis and Katzenbach 2021; Germundsson 2022; Hockenhull and Cohn 2021; Rieder 2018; Wijermars and Makhortykh 2022). Imaginaries of datafication differ within fields and depend on a variety of factors, such as organization and profession and historical immersion into data practices (Christin 2017). The realization of STIs by producing and implementing specific technologies depends on the social significance afforded the technologies in the context of where they are produced and embedded (Wajcman 2015). What is needed, according to Beer (2016), are close examinations of the discursive framings of datafication, giving accounts of how data-driven public administration has become embedded into organizational, political, and social life to reveal the dynamics of how it is sustained and gains power. "The power of Big Data [...] is in how those data and their potential are imagined and envisioned" (Beer 2016:9).

STIs describe both what is attainable and what ought to be attainable according to actors and institutions. Imaginaries shape and constrain technological practices (Groves et al. 2016). This is also about understanding the relationship between STIs and power dynamics through highlighting how some STIs achieve prominence while others are silenced or remain marginal (Ruppert 2019). Silenced issues have been of great importance to this research project as they point toward how datafication might be framed and embedded differently. In this dissertation, I

have continuously researched what public sector employees, organizations, and policy makers do in relation to datafication, how they translate, embed, and imagine datafication in their everyday practices. Following this, I also want to argue that a situated practice approach to datafication needs to include a research focus on what public sector employees, organizations, and policy do not do or omit in relation to datafication. Unseen issues or silenced topics and actors can point us to how datafication might be constructed differently. This argument is inspired by Carol Bacchi's (2009, 2012) "what is the problem represented to be" approach to policy analysis, in which she argues that we need to ask what is left unproblematic and if the problem can be thought about differently. By focusing on silences, both in practice and in the STIs, we are able to explore the critical potential of the situated practice approach.

It is then also of importance to understand how STIs enter, are translated, and are negotiated into assemblages of materiality according to Jasanoff and Kim (2015). Here, we return to the data assemblage. Although there is a variety of critical scholarship that points to the negative impact of datafication on society, this body of literature seems to assume that the STIs of the data-driven society materialize without interruption (Caplan et al. 2020). I challenge this assumption with the empirical cases presented in this dissertation. To align and assemble the apparatuses of the data assemblage takes practical work. This work is often challenged by already-existing practices as data assemblages in public administration often have a long history. Again, it is necessary to direct scholarly attention to the context of datafication and the importance of situating datafication within public administration and existing data practices (Dencik 2019). This then allows me to position myself as suspicious toward any notion of linearity connected to the way datafication changes public administration (Couldry 2004).

3.5 Summary and research questions

In the beginning of this section, I introduced four puzzles that the scholarly debate on datafication often encounters and that confound me in my quest to make sense of the empirical observations in the field. By laying out this theoretical framework, I aim to position my dissertation at the intersection of these puzzles. While each of the papers makes use of a distinct analytical framework to answer the individual research questions posed (see Table 2), there are some overall theoretical assumptions that tie all four papers together. I position my work at the intersection of practice theory and STS and operationalize this line of thinking by making use of two key concepts: STIs and the apparatuses of the data assemblage. By asking the overall research question of how public administration datafication is framed and embedded in Norway, I situate datafication in public administration. The work presented in all four papers is built on the simple assumption that to understand how datafication impacts society and how to respond to datafication, we need to understand how it is framed and embedded into the everyday practices of public administration.

I regard a situated practice approach as introduced by Dencik (2029) as fundamental to the actionable critique, which I will elaborate upon in Chapter 5 of this dissertation. While some scholars associated with critical data and algorithm studies have started to investigate the technological, social, and political and economic apparatuses of the data assemblage (Allhutter et al. 2020; Redden et al. 2020), I aimed to contribute to this literature by especially paying attention to the unseen issues/silences in STIs and practitioners' discourse and to the obstacles/perceived challenges when producing and aligning data assemblages in the public sector. These questions are especially important when working on Dencik's (2019) suggestion to uncover the potential for resistance or contestation and in my quest to work in close collaboration with public administration (Lassiter and Eric 2005). I therefore introduce two sub-questions to answer the overall research question of how public administration datafication is framed and embedded in Norway. These are as follows: Q1) What interests and agendas are inscribed into the STI of datafication, and what is omitted? Q2) How has this imaginary been constructed and embedded into the Norwegian public sector? Q3) What obstacles do practitioners experience when attempting to stabilize STIs into working data assemblages? In the next chapter, I will elaborate further on how this theoretical framework has been operationalized.

Table 2. Overview papers and research questions

	Paper 1	**Paper 2**	**Paper 3**	**Paper 4**
Research question	How does the context of public administration constrain the STIs of datafication as these are realized in specific data assemblages?	What challenges are encountered and problematized in a nascent phase of data-driven public administration implementation?	How are citizen perspectives problematized and included in policy and practitioner discourse in the datafication of public administration?	What are the policy problems for which these technological developments seem to be the best solutions, and how has the STI of data as a resource emerged in public administration?
Key concepts/ analytical framework	Sociotechnical imaginaries and data assemblage	Implementation theory	Ladder of smart citizen participation	Sociotechnical imaginaries and problematization
Focus	Embedding datafication	Framing and embedding of datafication	Framing and embedding of datafication	Framing of datafication
Sub-question	Q3	Q1 and Q3	Q1 and Q2	Q1 and Q2
Contribution to actionable critique	Tensions/obstacles in practice of doing datafication	Tension/obstacles and unseen issues in framing and producing datafication	Focus on one specific unseen issue	Unseen issues in framing datafication

4 Methodology

This research project made use of a multi-method approach in its quest to answer the overall research question of how the practice of datafication is framed and embedded in Norwegian public administration. The methodological choices made are briefly described in each of the papers. However, the limited word count of academic journal articles does not allow for comprehensive accounts of choices made in the method sections of papers. This chapter will provide an overall introduction to the research design of this dissertation and its origins, as well as a detailed account of method choices, in addition to reflecting upon the quality of the overall scientific project.

4.1 Research design: Investigating the data-driven public sector

While there is a variety of studies on the uses of data-driven systems and their implications, less prominent are studies on the production, construction, and framing of datafication. As argued in Chapter 3, this research project is interested in what people and institutions do in relation to datafication. The opacity of data-driven technology, tied to intentional secrecy, technical complexity, and the lack of specialized knowledge, has been problematized by a variety of scholars and has resulted in debate on the methodological approaches to this phenomenon (Bucher 2012; Diakopoulos 2014; Lomborg et al. 2020; Seaver 2017). Opacity can be less problematic in public administration as the Norwegian public sector is required by law to provide researchers access to its inner workings. Indeed, one of the key motivations to study the public sector was the accessibility of research sites. In addition, the focus on datafication practices rather than a single technology has in many ways circumvented black box issues.

In his paper "Thinking Critically About Researching Algorithms," Kitchin (2017) outlines, among other methodological approaches, an introduction on how to unpack the sociotechnical assemblage of data and algorithms. The research design of this dissertation has its origins in Kitchin's paper and approach. Kitchin argued to examine the full sociotechnical assemblage by including how these systems are framed and conditioned by forms of knowledge, legalities, governmentalities, institutions, and finances, in addition to investigating why data-driven systems are installed in the first place. Interviews and ethnographies of data teams and the institutional apparatus surrounding these are central. This can and must be supplied with discursive analysis of internal documents, legal frameworks, and policy documents to gain a better understanding of the practices and structures of data teams and institutions and to reveal how datafication is imagined, promoted, and legitimized according to Kitchin. By taking an ethnographic approach to these systems, we see features that are elided or obscured (Kitchin and Lauriault 2014). This is also in line with Dencik's (2019) suggestion to situate data practices in relation to other social practices within a specific context. The value of studying the practices of embedding and producing data-driven practices is demonstrated, for example, by the work of Allhutter et al.

(2020). Researching the algorithmic profiling of job seekers, the authors investigated the inherent politics of data-driven technology through an in-depth analysis of technical documentation and policy documents to understand the conceptual, technical, and social implications of this specific system. Additionally, Redden et al.'s (2020) study of datafied child welfare services and work on unpacking data assemblages by showing how systems of thought, ownership structures, legal frameworks, and organizational practices influence predictive data systems has been of great inspiration here. More recently, Chaudhuri (2022) ethnographically unpacked the underlying assemblages of a public welfare distribution system in India and shows how algorithmic sorting is enacted in relation to various human interactions, institutional contexts, and databases, drawing on Kitchin's (2017) work.

This dissertation then presents an explorative case study on the inner workings of public administration datafication. Following Timmermans and Tavory (2022:9), the project does attempt to focus and defocus its research lens by zooming in on specific questions and zooming out at the interconnectedness of social life, which required a multi-method approach. As argued in the previous chapter, at its core, this research project can be placed within social constructivist studies. Its ontological and epistemological routes are within STS and practice theory, and the research project primarily makes use of qualitative research methods. STS has placed significant emphasis on the importance of ethnographic field studies on scientific practices and technology production and the unpacking of black boxes, analyzing both production and product (Latour 1987; Star 1988), which have been of great inspiration to this project. Furthermore, public administration research encourages empirical and qualitative work to understand processes within the public sector (Yang and Miller 2008; Haverland and Yanow 2012).

As this is the first empirical account of public administration datafication in Norway, the demarcation of the project has been highly challenging, and an explorative approach seemed most fitting. There was, and still is, an extensive research gap in the Norwegian context and the scholarly debate on this topic. In 2017, when the project was developed, there were very few researchers studying data-driven systems in public administration as the main focus of the field remained in accounts of big tech's role in datafication. Abductive analysis, in which an element of surprise steers the research process, has not only been a key element of the analysis in the papers presented but also of the overall research design of the project. Beginning with in-depth field work of a specific data team in the Labor and Welfare Administration (Reutter 2018), there were three moments of surprise that steered much of the further development of the research project: First, my observations seemed to correspond very little with the research done on internet platforms that I had read earlier; the data team observed seemed to struggle greatly in their work. There was a significant misalignment observable between the imaginaries identified in previous research - where the public sector can make use of fast and cheap data analysis, glean data from entire populations, let data speak for themselves and achieve an impartial view from nowhere under datafication - and my own observations (Rieder & Simon 2016). Second, AI seemed to

present just one of many elements in an overarching focus on data recirculation/use in public administration. Third, policy documents and political decisions appeared to play a key role in what was happening in the data teams. This suggested to me to broaden the perspective on what was researched (see Chapter 2: Clarification of Concepts) and to include different methodological approaches in the project, following Timmermans and Tavory's (2022) suggestion of agile adjustments to the research project and continuously defocusing and refocusing the research lens. Simply focusing on two isolated data teams in the public sector could not help me understand these moments of surprise. The presented papers thus focus on different aspects of datafication and provide insights into both specific projects and the whole of the public sector.

Paper 1 is based on field work conducted in two data teams and investigates how the sociotechnical imaginaries of datafication are materialized in concrete settings within the public sector. It focuses on the multiple obstacles and contestations that data team members face when producing data-driven technology. However, while conducting field work in the data teams, I also realized that what I observed was about more than AI and big data in the specific project observed. Zooming out, I then provide an overview of activities on datafication in the Norwegian public sector in Paper 2. This paper addresses the perceived challenges in the field to realize the grand idea of data-driven public administration. It aims to demystify data-driven public administration, in addition to raising the issue of datafication within public administration studies. This is, to our knowledge, one of the first paper published in a public administration journal mentioning the term "datafication". In this paper, we also identify several unseen issues, such as the lack of citizen involvement in the discourse on data-driven public administration. The third paper builds on the unseen issues and makes use of much of the same data to understand how citizen perspectives are problematized among practitioners and in policy papers. We also included new data sources, such as the method sections of each policy paper and web page material describing how the policy papers were produced and who was consulted. As argued in Papers 1 and 2, the public sector faces a variety of challenges when attempting to materialize the STIs of datafication. Despite the hype of AI and big data, little seemed to have been achieved. Paper 4 then fully focuses on policy papers as these are identified as key sites for maintaining and embedding the STIs of datafication. The research design is summarized in Table 3.

In my research, I have been inspired by the idea of collaborative ethnography, which is attentive to public issues and collaborates closely with the general public and affected communities (Lassiter and Eric 2005; May and Pattillo-McCoy 2000). The researcher here invites commentary from the researched and reintegrates feedback into the research process. In this case, I had an ambition to not only research public administration datafication from an outside perspective or participant observation but also to actively discuss my findings with the public sector. Although I have not generated the research questions themselves in collaboration with the public sector, my observations have been continuously fed back to practitioners in the Norwegian public sector. My observations and possible explanations of what was observed have been discussed with people

working in and through the data-driven public sector throughout the research project. While collaborative ethnography is often concerned with producing an ethnographic text itself with consultants in the field (Lassiter 2005), I have chosen more informal modes for practitioners to review and comment on my findings. This includes, for example, meetings with data team members about obstacles in connection to Paper 1, where I present a drawn model of the variety of obstacles, or the presentation of preliminary findings and unseen issues at public sector conferences in connection to Papers 2, 3, and 4. Presenting my work, and then discussing it with practitioners at conferences, has proven especially valuable, and a list of these interactions is attached in Appendix B. Although the informal discussions with practitioners at all levels of the public sector are not formalized in a research paper, I regard them as important elements of this research project. These conversations have, for example, led me to the idea of the actionable critique further elaborated on in the final section of the dissertation, as well as sensitivity in how I present my work to different audiences. In addition, my presence in the field has been eased by access to public administration as many practitioners working on datafication have met me or seen me speak at internal conferences.

Table 3. Research design for Papers 1-4

	Paper 1	**Paper 2**	**Paper 3**	**Paper 4**
Title	Constraining Context: Situating Datafication in Public Administration	Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation	In Search of the Citizen in the Datafication of Public Administration	Public sector data as a resource: tracing the emergence and embedding of a sociotechnical imaginary
Research Question	How does the context of public administration constrain the STIs of datafication as these are realized in specific data assemblages?	What challenges are encountered and problematized in a nascent phase of data-driven public administration implementation?	How are citizen perspectives problematized and included in policy and practitioner discourse in the datafication of public administration?	What are the policy problems for which these technological developments seem to be the best solutions, and how has the STI of data as a resource emerged in public administration?
Methodology	Qualitative approach	Qualitative and quantitative approach	Qualitative and quantitative approach	Qualitative approach
Method of data collection	Field work, semi-structured interviews, meeting observations, and document analysis	Online survey (n=35), semi-structured interviews, and document analysis	Online survey (n=35), semi-structured interviews, and document analysis, web page information on policy production	Publicly available policy documents
Scope	Data teams in two public sector organizations	Meso-level, public-sector organizations	Meso-level, policy on digitalization and public sector organizations	Policy on digitalization, national efforts
Method of analysis	Abductive analysis	Abductive analysis	Abductive analysis	Practice-oriented approach to document analysis

4.2 Research site and scope: The Norwegian public sector

Before digging deeper into how data were collected, it is important to briefly reflect on the research site and scope of the project. This dissertation takes the Norwegian public sector as its case, which clearly distinguishes my research from single project accounts of data-driven technology (see, for example, Allhutter et al. 2020) or single public sector area accounts (see, for example, Germundsson 2022 on social welfare or Nikunen and Hokka 2020 on public broadcasting). The scope of the research project was continuously adapted throughout the project, zooming in and out and allowing me to research datafication as a "continuously constructed project, shaped by multiple, converging and conflicting forces" (Dencik, 2019:246), across several sites and on different administrative levels. Each of the papers includes a short section on the Norwegian public sector to contextualize the presented findings and thus allow other scholars to evaluate the relevance of the findings for their own research. These sections, for example, address the importance of register data in the Nordics (Frank 2000) and discuss specific Norwegian practices, such as corporative pluralism (Arter 2004; Rokkan 1966) and the use of hybrid advisory committees in policy making (Christiansen et al. 2010; Krick and Holst 2021).

What, then, characterizes this case? First, the Norwegian welfare state has been established as a social democratic welfare state, with high levels of decommodification and a focus on universal provision of social security to all citizens through the state and the use of taxes to do so according to Esping-Andersen (1990). Here, the core values of solidarity, universal rights, equality, and a participatory democracy were central guiding principles. As with many other welfare states, the Norwegian public sector has undergone significant changes since its establishment through, for example, the introduction of new public management (Christensen et al. 2008). The welfare state does, however, remain a key actor of the Nordic model (Dølvik 2013), which also requires a well-functioning public administration. Norwegian public administration is heavily regulated by law. The most important law for this dissertation is the Public Administration Act that among other regulations defines citizens' rights in relation to the public sector, such as the right to explainability and transparency in case work, which are also key element in discussions on the automation of public administration (Schartum 2017). It is important to note that Norway currently is not part of the European Union. Norway has, however, signed the European Economic Area Agreement (EEA) and is therefore heavily influenced by EU law-making.

Today, the Norwegian public sector consists of a variety of administrative levels and public sector organizations; it is both large and fragmented. On its most basic level, Norwegian public administration can be divided into state agencies and local government agencies (Fimreite and Grindheim 2007). In this dissertation, I have mainly focused on state agencies in researching the datafication of public administration rather than local councils as done, for example, by the Data Scores as Governance project (Dencik et al. 2019; Redden et al. 2019). State agencies such as the Norwegian Labor and Welfare Administration and the Norwegian Tax Administration operate on a national and centralized basis, with local offices managed by municipalities. These agencies are

especially interesting as they operate at the intersection between the political system, administration, and society, developing technology and strategies that are then introduced to local contexts (Fimreite and Grindheim, 2007). I have thus chosen to focus on national efforts rather than municipal efforts of datafication. While Papers 2, 3, and 4 focus on the whole of the public sector, Paper 1 presents a case study on two specific data teams within two Norwegian public agencies. It is important to note that intelligence agencies and the military, although part of the public sector, are not included in the analysis. This is because these sections of the public sector are less accessible for researchers and often require security assessments. As these sections of the public sector are also not addressed in national datafication policy, I want to argue that my research still largely represents the public sector on a system level.

4.3 Data generation

The project began with field work in the Labor and Welfare Administration (NAV) that was later supplemented by field work in the Norwegian Tax Administration (SKATT). The field work included the close shadowing of specific projects aimed to investigate how two data teams produced data-driven technology and how they embed datafication into everyday work practices in each of the organizations. The survey and follow-up interviews were intended to provide a broad overview of activity in Norwegian public administration and gain a better understanding of the various activities and practices across organizations. This also provided a basis for investigating who was problematized and included in framing and doing datafication in Norway. The motives, motivation, and imaginaries of this activity were then researched through a document study of policy papers. This has by no means been a linear research process; most of the activity was conducted simultaneously, and the various methods of data generation have informed and inspired each other.

4.3.1 Field work, meeting observations, and interviews

Neyland (2016), Seaver (2017), and Kitchin (2017) encouraged ethnographic work when researching how data-driven systems are produced and embedded into practice. Especially in the beginning of this research project, qualitative ethnographic research among practitioners doing datafication was rare to non-existent. In addition, much existing empirical work in critical data and algorithm studies seemed to concentrate on already-working systems from an outside perspective rather than researching datafication in the making. When I found an online article announcing the establishment of an AI lab in the Norwegian Labor and Welfare Administration in 2017, my scholarly curiosity was sparked, and I decided to attempt to study the work done by this team. Ethnographic work on the production of data-driven systems has since proven highly valuable to critical data and algorithm studies (Chaudhuri 2022; Passi and Sengers 2020).

Access to NAV was gained in December 2017 in the course of the work on my master's thesis (Reutter 2018). The team had started working on AI some months previously, and there were, at that point, just three core members and a team lead working on data-driven technology in the

organization. However, this team gained members rapidly over the course of the field work, and the IT department underwent several re-organizations. The interviews, field notes, and meeting observations from the first field work at NAV in connection to my master's thesis were included in the analysis of my PhD project. This has then served as pilot field work for this dissertation. I kept in touch with the data team and leadership after the initial study and re-established contact when starting the presented research project in autumn 2018. When I restarted field work with the NAV data team, I followed them to several conferences and interorganizational meetings. This included meetings with the AI fagforum, a public sector meeting place where people working on data-driven technology (both technical and organizational) were meeting to discuss projects and exchange thoughts and experiences. In one of the first meetings, I established contact with the Norwegian Tax Administration (SKATT), which, at that time, was seen as a forerunner in data-driven technology in Norway. NAV and SKATT represent in many ways the heart of the Norwegian welfare state as they are the two biggest organizations in the public sector, with responsibilities touching upon every citizen's life. They were also chosen as cases as neither rely on private sector consultancy firms in their datafication efforts, which again has eased access to the field. The particularities of each data team are described in the method section of Paper 1.

How, then, have I legitimized my presence in these organizations? As argued previously, I have been an active collaborator with the public sector (May and Pattillo-McCoy 2000). In the initial meetings with the data teams, I presented a short literature review on datafication in society and critical scholars' work in this field more generally (including work by Boyd and Crawford 2012; Pasquale 2015; Ziewitz 2016; and Zuboff 2019). I stressed that there was very little knowledge on how big data and automated decision-making is produced in public administration in the social sciences and introduced them to Kitchin's (2017) idea of unpacking the socio-technical assemblage of data and technology, as well as my intentions of conducting field work in the data team. Both teams welcomed research on their inner workings, although I was clear about my ambitions to not only map practices but also critically interrogate datafication. I think I was also lucky as I was the first social scientist showing interest in issues of datafication and the data teams seemed therefore also curious about my work and perspectives. Prior to the field work, we discussed the practicalities of field work (key card, office space, etc.) and any concerns the participants had. A contact person in each team helped me organize interviews and meetings and kept me informed on the data teams' work routines.

Practically, I spent the field days in both organizations sitting in the open landscape, following data team members to meetings, and scheduling one-on-one interviews and meeting with specific participants both inside and outside of the data teams. Most notes were taken on a small physical notebook as I could easily take the notebook with me everywhere and write down any observations. The field notes taken were directly transcribed into NVIVO, which I used as a searchable data storage for my empirical data material as NVIVO allows a variety of different file

formats. At the end of every field day, I wrote a short reflection note about my own role as researcher. An overview of activities is presented in Table 4.

Table 4. Overview of field work at Norwegian Labor and Welfare Administration and Norwegian Tax Administration

Norwegian Labor and Welfare Administration	
Field work 1 January 2018	13 days of field notes 6 meeting observations 11 interviews 9 internal documents and PowerPoint presentations
Field work 2 March 2019	5 days of field notes 9 meeting observations 6 internal documents
Field work 3 April 2019	3 days of field notes 3 meeting observations 4 interviews 3 internal documents
Other November 2020	2 follow-up interviews on progression of sickness-benefit project
Norwegian Tax Administration	
Field work 1 September 2019	3 days of field notes 2 meeting observations 5 interviews
Field work 2 October 2019	5 days of field notes 2 meeting observations 5 interviews 3 PowerPoint presentations
Field work 3 November 2019	5 days of field notes 2 meeting observation 3 interviews
Others	
	3 days of field notes from AI fagforum observation 2 days of field notes from Digitaliseringskonferansen 19/20 4 days of field notes from NOKIOS 18/19 Letters of allocation from the Ministry to both organizations, 2016 to 2020 Virksomhetsstrategi SKATT

I had an ambition to study the practice of datafication in the public sector. It was therefore crucial to conduct the field work in several stages. As suggested by Tavory and Timmermans (2022), I entered and re-entered the field several times to familiarize myself and de-familiarize myself from the empirical work. Instead of conducting the field work in one take, I conducted several short stays with the data teams. The time between these field days was spent reading, analyzing, and making hypotheses that I then tested in the field or discussed with participants. This research design allowed me to creatively engage with the data teams and the surrounding organization, as well as engage in other data generation work (surveys, interviews, document studies) on other research sites. During the field work, the research approach needed several adjustments. In my first week at NAV, I struggled with the interview flow, but I noticed that many of the data team

members and public sector workers were using PowerPoint in their meetings. I decided to encourage my participants to bring PowerPoint presentations about their work to interviews as a conversation starter. In addition, I started calling interviews "meetings." This contributed significantly to ensuring a professional yet relaxed atmosphere as meetings are an important part of the day-to-day work practices in public administration. After Field Work 1 at NAV, I also stopped recording interviews on a recording device and only used written notes for data generation. Although this means that not every word was written down, I am confident that it also eased access to participants as it replicated a workplace setting. I took rigorous notes of the interviews and transcribed these directly after each meeting. In addition, I kept all the original notes and digitized them to keep together with the transcribed meeting notes.

4.3.2 Survey and follow-up interviews

As Redden (2018) and others have pointed out, there is often a lack of overview on the various activities of datafication. In addition, as highlighted previously, I soon realized that the practices observed in the data teams presented just one of many elements in an overarching focus on data recirculation/use in public administration. The question "Are there other places you should be?" (Timmermans and Tavory 2022:21) was answered affirmatively. The research lens needed adjustment to focus on the public sector more generally. Therefore, the second part of the data generation made use of a survey distributed through the AI fagforum, an interorganizational forum for data workers across the Norwegian public sector. This mapping was closely connected and legitimized through the government work on a national artificial intelligence strategy in the summer and autumn of 2019. It was intended to produce a report that could be used by policymakers in the production of a national strategy,[2] in addition to providing us with scholarly insights into the inner workings of public administration datafication.

The AI fagforum meets regularly to present and discuss issues of machine learning and other associated technologies and had around 40-50 member organizations in 2019. Through the field work described in Section 4.3.1, I became a member of this forum. Heather Broomfield, my co-author, was part of the organizing team of the forum at that time. This, therefore, seemed like a well-fitted arena to distribute our survey and access the people working on and through datafication in the Norwegian public sector. However, that also means that our data collection was limited to organizations that had already started or were planning to work on data-driven technology. Although surveys are not usually used in STS studies, they are a key tool within public administration studies to gain oversight of activities (Liu and Yuan 2015; Moon 2002). We used Redden's (2018) work on the Canadian public sector as a template to map the challenges of becoming data-driven on a Likert scale across the Norwegian public sector, in addition to free text opportunities. The survey template is attached as an appendix to Paper 2 and 3. As pointed out previously, we were especially interested in how datafication aligns and misaligns with existing

[2] The report has been sent to the Norwegian Ministry of Local Government and Modernization and is publicly available at https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2634733

public administration practices to uncover tension in the practice of assembling new data practices. The survey was answered by practitioners (n=35) in 26 public organizations. Further details on the survey can be found in the method section of Papers 2 and 3. The sample size is of course small and therefore unfit for quantitative analysis beyond descriptive overviews.

We did here decide to make use of a two-phase explanatory research design (Cresswell and Clark, 2011). Follow-up interviews (n=12) were conducted to nuance the overall observations in the survey. In these interviews, we met the survey participants to discuss their survey answers and provide them with the opportunity to reflect more generally on their work practices. We began each interview with the question "What does it mean to you and your organization to become data-driven?" To facilitate an open dialogue, we visualized their survey answers on perceived challenges and asked them to reflect on the answers in the semi-structured interviews. The interview guides are attached to Paper 2. Again, as with the field work conducted, participants were allowed to bring PowerPoint presentations or demonstrate projects on screen. Interviews lasted between 60 and 90 minutes. The interviews were recorded, transcribed, and anonymized. The practitioners surveyed and interviewed were system-level designers and not street-level bureaucrats, as most activity was observed in this administrative level at that time. We then supplemented the survey and interview data with policy and strategy documents to be able to study the data-driven public administration at the system, organizational, and individual levels (Pencheva et al. 2020).

4.3.3 Document study

Document studies have been used in all four papers to some degree as I regard them as a key source for understanding how the datafied public sector is framed and embedded in Norway. Paper 4 stands out as it utilized document analysis as its only research method. In this dissertation, I have followed Asdal and Reinertsens's (2020) practice approach to document analysis, acknowledging that documents bring issues into being, mobilize resources, promote actions, and facilitate opportunities. Jasanoff and Kim (2009) argued that national policies on technology are useful sites to research STIs and the role of political practices in stabilizing and mobilizing these imaginaries. Researchers such as Allhutter et al. (2020) and Redden et al. (2020) included policy documents in their analyses of public administration datafication.

How than have the analyzed documents be selected? A quick overview of key policy documents in Norwegian public administration datafication can be found in Fig. 1. An appendix summarizing each of the policy documents has in addition been attached to Paper 4. I regard these documents as constituting the core of Norwegian datafication efforts, as they all discuss different aspects of data sharing and use within and outside the public sector. In addition, these documents address the whole of the public sector, rather than specific subject areas such as policing or child welfare. Three out of the six documents were published after this research project was initiated, which again indicates the importance given to issues of datafication in Norwegian public administration.

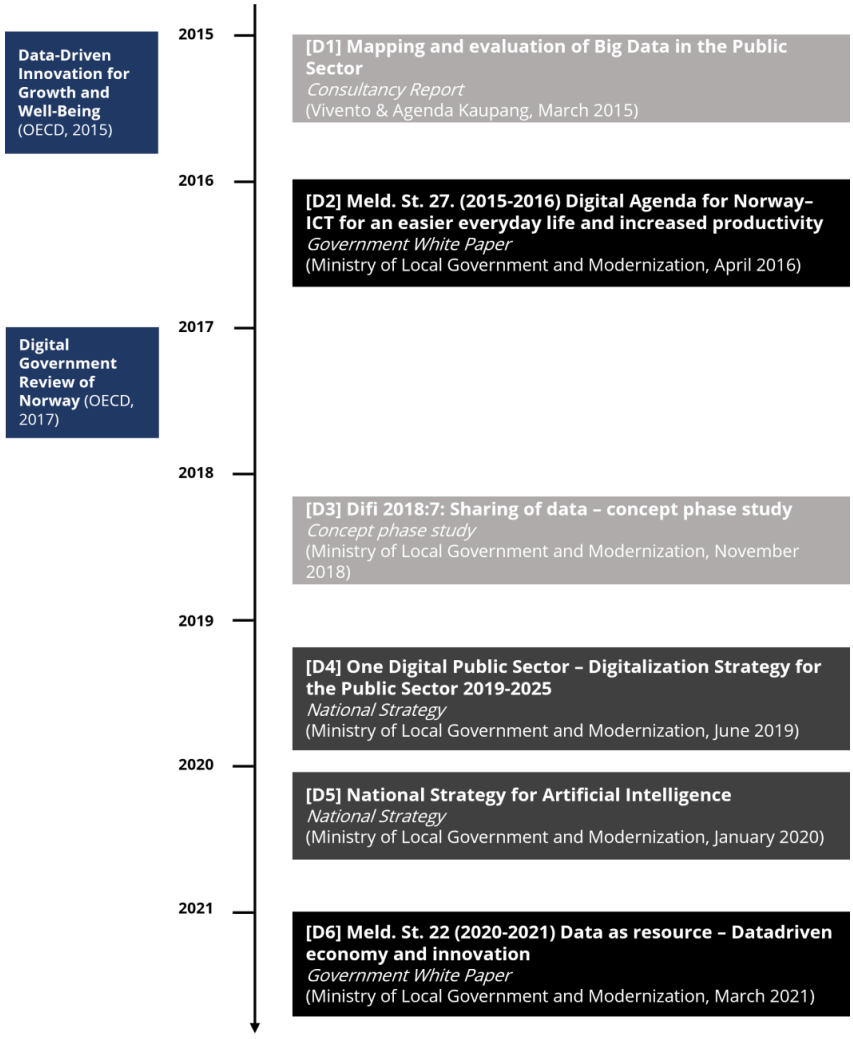


Figure 1 Network of datafication documents

Although policy and strategy documents were the key documents analyzed during the research, other documents have been part of the data material as well. This includes web page material analyzed in Paper 2 and internal strategy documents and PowerPoint presentations analyzed in Paper 1. Including PowerPoint presentations in the analysis was part of taking the context of public administration seriously, as PowerPoint presentations are a central element of communication and a central work tool for data workers. These secondary documents were then helping me to contextualize my research findings. I will elaborate on how documents and other data material was analyzed further.

4.4 Data analysis

Abductive analysis has been central to the overall analytical approach of this dissertation, where the element of surprise has sparked ideas and resulted in the theory-matching process (Tavory and Timmermans 2014; Timmermans and Tavory 2022). Abductive analysis is especially well suited as the research was characterized by an explorative approach and an open research attitude, allowing me to draw inspiration from various disciplines, fields of study, and theoretical perspectives (Charmaz 2006; Reichertz 2010). This is also highly relevant as critical data and algorithm studies themselves are an interdisciplinary collection of work combining a variety of theoretical traditions. While abductive analysis provided the backdrop for the research project, other more practically oriented analysis approaches supplemented abduction in each of the papers.

Abductive analysis is about using elements of surprise as a core tool in the research process (Tavory and Timmermans 2014). Research astonishment arises in relation to existing knowledge, theory, and beliefs, and, therefore, it requires us to question our own knowledge and to scrutinize the data material (Charmaz 2006). Surprise is a highly subjective perception, as Rinehart (2020:305) stresses: "What constitutes a surprise will not necessarily be agreed between researchers and across groups concerned." What I find surprising in my research will therefore differ from other researchers. The central puzzles presented in the theory chapter previously are more than surprises: they are irritations that arose when dealing with the messy empirical world. "[…] it refuses to be ignored. It niggles, it irritates" (Rinehart 2020:305). There were three central irritations that steered this research project. First, my observations seemed to correspond very little to the research done on internet platforms that I had read earlier as the practitioners struggled to make systems work. Second, AI seemed to present just one of many elements in an overarching focus on data recirculation and use in public administration. Third, policy documents and political decisions appeared to play a key role in what was happening in the data teams. The data teams were highly restricted by practices outside of their own realm. I will elaborate on how this has affected and steered the analysis process further.

According to Timmermans and Tavory (2022), the abductive research process consists of several interlinked and often non-linear stages. This includes familiarization with a variety of theories,

facilitation of elements of surprise, and coding and re-coding of data. I began with field work at NAV in 2018 and 2019. At that point, I was a relatively inexperienced researcher dealing with an under-researched field. As a media and STS scholar, I was familiar with a variety of literature on algorithms, data, and the practice of assembling technology. When entering the field, I did not have a clear research question, but I did have the explorative ambition of observing the practice of how an AI system is assembled. As pointed out earlier, this research focus soon appeared to be too narrow. Hence, there was a "nagging" experience of the need to adjust the research scope as AI seemed to present just one of many elements in an overarching focus on data recirculation and use in public administration. This led me to the data collection conducted in Paper 2, in which the whole of the public sector and the idea of becoming data-driven stand central. Furthermore, what was observed did not correspond with the literature I had engaged with earlier, and the data team members struggled to make data-driven technology work. I therefore began to actively engage with a variety of obstacles and challenges that practitioners and data team members experienced in both the collection and analysis of data. These then became central elements of Papers 1 and 2, and indeed the whole of the dissertation. Paper 1 had its very origin in a single meeting when members of the NAV and SKATT teams met up to discuss their work routines. One of the team leads drew up a simple data science model used by Microsoft and then started to alter this model to fit it into the public administration context together with me and the other meeting attendees. The issues discussed in this meeting then resulted in the analysis presented in Paper 1.

Paper 2 zoomed out and tried to understand how the production of machine learning is part of a bigger change in public administration. After collecting the survey data for Paper 2 and performing a first reading of interviews, we were struck by the immense emphasis on organizational issues in datafication processes. There was, however, little focus on these issues in the critical data and algorithm literature we had consulted earlier or in the policy documents with which we were familiar. This mismatch between policy prioritizations and practitioners' discourse, in addition to the seeming absence of several issues discussed in critical literature on public administration datafication, provided an element of surprise that was further explored in the discussion of Paper 2. We analyzed the interviews inductively according to meaning, focusing especially on the many challenges drawn up and discussed. A policy analysis in which we focused on conceptualizations of datafication, and the prioritization of policy action then supplemented this analysis. This allowed us to find both discrepancies and similarities between practitioners and policy makers' ideas about data-driven public administration, in addition to being able to identify unseen issues often drawn up in critical literature.

After an initial reading of the interview data and a consultation of policy documents, another issue arose that sparked our curiosity: where is the citizen in all this? Citizen perspectives seem to be completely lacking in any of the data material collected. This was surprising given both the democratic focus on including trade unions and civil society in decision-making in Norway (Krick

and Holst 2021; Rokkan 1966) and an empirically observable notion of user-centered services. Our curiosity was sparked, and we re-entered the field in search of the citizen in public administration datafication. The question then was whether citizens are able to challenge and alter how datafication is embedded and framed in the Norwegian public sector. We started by reading and re-reading both interviews and policy documents searching for any traces of citizens/users/customers in context to be able to investigate how citizens were problematized and envisioned. In addition, we included accounts of how policy documents were produced and who was consulted in this analysis as this information is publicly available. By using and adapting Cardullo and Kitchin's (2019) framework of smart citizen participation, we were able to investigate this further, focusing especially on the two central concepts of user-centric and needs-based approaches drawn up in both interviews and policy.

Simultaneously while writing Papers 1, 2, and 3, there was another frustration bugging me: after continuing my field work at NAV and SKATT and focusing on obstacles and challenges in two of the papers, I was again and again struck by a simple question: why is public administration trying so hard to make datafication work when there are a lot of projects failing? What drives datafication? In other words, although datafication was dominating in policy and discourse, I struggled to find working systems. I felt a need to go back to investigate how datafication was legitimized in policy documents, the STIs and their role in embedding datafication in public administration, which resulted in Paper 4. This paper is based on an in-depth document analysis of policy papers. While Papers 1, 2, and 3 use documents as a secondary data source, Paper 4 puts the documents at the center. Documents seemed to present the most concrete and materialized version of STIs. Combining Carol Bacchi's (2009, 2012) "What is the problem represented to be?" (WPR) approach to policy analysis and Asdal and Reinertsen's practice approach, we analyzed the documents asking the following open empirical questions: What happens in this document? For whom is this document a tool, and who are the addressees? What is the policy problem represented to be? How has this representation of the problem come about, and which actors speak on behalf of datafication? How does the text establish authority—that is, which justifications and arguments are mobilized? What is left unseen or even omitted?

Again, abductive analysis is about using elements of surprise, which are highly subjective and affected by theories and perspectives that a researcher is familiar with as a core tool in the research process (Tavory and Timmermans 2014). Other valuable lines of inquiry could have been taken. As Timmermans and Tavory (2022:160) argue, "Surprises provide you with a space to broach a topic and ideally open up the dialogue for others to listen to what you are about to say." This, I hope I have achieved through the four papers presented in the dissertation.

4.5 Situated knowledge production quality of research and research ethics

In the social sciences, the most common way of evaluating and reflecting on the quality of empirical research is through the positivistic notions of reliability, validity, and generalization, which have their origins in the natural sciences. These criteria are, however, of little relevance for qualitative studies. Tracy (2010) therefore introduces eight criteria for excellent qualitative research that can be achieved through various means, practices, and methods. These criteria include a worthy topic, rich rigor, sincerity, credibility, resonance, a significant contribution, and ethical and meaningful coherence. Rather than discussing each of these criteria in relation to my dissertation, in this last section, I want to shortly reflect on my own situatedness, some core limitations, and the ethical aspects of researching the data-driven public sector.

In this dissertation I have followed the constructivist methodological approach, recognizing the important role of the observer and society in constructing the patterns we study in the social sciences (Moses and Knutsen 2012:9). As Law (2004) argued, we here replace an innocent, neutral, singular, and technical understanding of method. Method produces realities; it re-works, re-bundles, and re-crafts realities and creates new versions of the world. Haraway (1988) furthermore argues that there is no such thing as a neutral standpoint or the possibility of detachment for the social scientist. Observation statements hence depend on the perspective of the investigator and the acknowledgment of situatedness. Using abductive analysis required me to familiarize myself with a variety of literature on data and algorithms and technology in society, as well as with Norwegian public administration, before entering the field (Tavory and Timmermans 2014). This then has influenced the questions asked and angles taken in the field. Prior knowledge about technology and public administration have however also been of great importance in engaging and collaborating with participants and institutions. There is often a great distance between qualitative researchers and people working on and through datafication (Moats and Seaver 2019). As Atkinson and Morriss (2017) argue, practical competence is a valuable tool in the analysis of specialized activity (such as public administration) and in engaging with the observed. What is required then is to maintain a reflexive awareness and an analytical distance to what is researched, which is also central in discussion of collaborative ethnography (Lassiter and Eric 2005)

In his review of how to unpack the sociotechnical assemblage of data and algorithms, Kitchin (2017:25) argues that being able to gather a vast amount of data and interlink them is no easy task, but it is manageable, especially if undertaken by a research team rather than a single individual. An obvious limitation of this dissertation is then that most of the data collection and analysis was undertaken by me or in collaboration with one other researcher. I did not have the resources or contacts to assemble a research team to conduct a large-scale case study on the Norwegian public sector. The survey used in Papers 2 and 3 was, for example, not able to reach all organizations in the Norwegian public sector, and the field work presented in Paper 1 only

followed the data team for a finite period. This dissertation, therefore, only presents a partial picture of public administration datafication in Norway at a specific point in time. Nevertheless, I did not abandon or exit the field completely after the survey, interviews, and field work, but I kept actively engaging with public administration until the very end of the research project to both confirm and discuss findings, which has had a positive effect on the overall quality of the research.

I have treated "access as a kind of texture, a resistance to knowledge that is omnipresent and not always the same" (Seaver, 2017), constantly negotiating with the practicing field and attempting to find new angles when necessary. Despite the difficulties other researchers have experienced when attempting to study the practice of doing machine learning (see, for example, Diakopoulos 2014; Neyland 2016), I experienced the public sector and especially the AI fagforum and the two field sites as open and welcoming to my research and the findings presented. The question then is to what degree the processes of datafication can practically be unpacked at all. Some aspects of datafication will always remain black-boxed or inaccessible to the researcher (Gillespie 2014b). I want to argue that decentering data and technology in researching datafication can bypass several of the black box problematics researchers face when approaching data-driven technology (Bucher 2012; Kitchin 2017). I did not, for example, have access to the code produced by the data teams in the field work presented in Paper 1. Nevertheless, I was able to present interesting insights into the inner workings of how datafication is embedded and produced in Norwegian public administration. Indeed, I did not feel a need to access all the technology or the data sets.

Finally, I want to shortly reflect on the ethical aspects of the data collection and analysis. The Norwegian National Committee for Research Ethics in the Social Sciences and the Humanities has laid out guidelines for research ethics, which I familiarized myself with before embarking on this research journey (NESH, 2022). As the Committee does, I regard one of the key quality criteria of qualitative research as transparency. Through laying out the research design and reflecting on the data generation and data analysis in this section, I have attempted to be transparent in the way the research was conducted. I have also consistently used extracts and examples from the original data material in papers to enhance transparency in my analytical endeavors.

For Papers 1, 2 and 3 I worked in close collaboration with public administration and the many people that make up the public sector. The overall legal aspects of the project were assessed by the Norwegian Center for Research Data (the assessment can be found in Appendix A). All data material was transcribed, anonymized, and stored on a safe and approved university server. Participation in research should always be based on information and consent according to the Committee. The public sector employees interviewed in Papers 2 and 3 gave their written consent to the study prior to interviews. In the field work, however, consent to participation was given by leadership on behalf of the participants. Interviews were always voluntary. Some of the data team members were not comfortable with one-on-one interviews, which both I and the leadership respected. I made sure to inform every individual participant about the aim and purpose of the study prior to interviews. In meetings with several participants, I was always introduced by one of

the participants as a researcher and often asked to say a few words about myself and my research. Meetings involved participants indirectly affected by the research, and I therefore decided against using recording devices, as I did not always have an overview of each of the participants present

The Norwegian public sector is by law required to provide access to its inner workings to researchers; this does not mean, however, that researchers can bypass "human dignity and [consideration for] [participants'] personal integrity, safety and well-being." (NESH, 2022). While we decided to anonymize the public administration organizations researched in Papers 2 and 3, I used a different approach in Paper 1. I decided together with the organizations researched in this paper that while all employees were anonymized in the paper, the organizations themselves should remain visible. Full anonymity could therefore not be granted. In the data collection for Papers 2 and 3, we wanted the public administration employees to reflect critically on the datafication efforts in their organizations in interviews and surveys, and we therefore used anonymity strategically. As Redden (2018) argued in her mapping of the Canadian public sector, critical voices within public administration often seem to be silenced. This, then, was a trade-off between anonymity and transparency. As I was able to build a stronger relationship with the participants researched in Paper 1, I did not regard a fear of voicing concerns as pressing. I signed a non-disclosure agreement with NAV. This was, however, not intended to limit or censor my research. As I was sitting in meetings and in the open landscape, there was a small risk of me being able to see sensitive information about citizens. It was therefore important that I complied with the confidentiality regulated by the Public Administration Act. SKATT did not require such an agreement; we did, however, write up a small collaboration note in which the terms of field work were formalized.

Overall, this research project presents a partial and situated account of how datafication is framed and embedded into Norwegian public administration. This situatedness is further complicated by my ambition to collaborate with the public sector, constantly feeding back research findings and discussing my thoughts with the participants. The dissertation is a first attempt to understand the inner workings of public administration datafication in the Norwegian context as informed by critical data and algorithm studies. Nevertheless, as I have shown in the preceding chapters, this dissertation presents both a relevant and timely topic and a significant contribution to the field; the research is conducted with rich rigor, sincerity, credibility, and ethical considerations, and it therefore offers, according to Tracy (2010), qualitative research of high quality. Partial insights and situated knowledge are valuable both politically and analytically (Haraway 1988) as demonstrated by the findings presented in the next sections.

5 Summary of papers and key findings

The research papers that make up the core of this dissertation present four distinct yet interlinked insights into the inner workings of public administration datafication. The papers were purposefully published in three different journals to reflect the interdisciplinary approach to the research question and to disseminate the findings among different groups of interest. In what follows, I will summarize each paper and present the key findings. This will constitute a point of departure for the discussion in the next chapter of the dissertation.

5.1 Paper 1: Constraining Context

Published in New Media & Society.

This paper was part of a special issue on algorithmic governance in context (Gritsenko et al. 2022) and had a clear ambition of situating datafication beyond big tech. While the inner workings of producing data-driven technology in the private sector often remain inaccessible, the ever-growing interest in data and algorithms in public administration presents a compelling field site to research how STIs of datafication are materialized in specific data assemblages. Two data teams at NAV and SKATT were followed in their everyday work practice of producing data-driven technology in 2018, 2019. The paper presents the first empirical account of the inner workings of datafication in the Norwegian public sector and offers a valuable contribution to critical data and algorithm studies in the Nordics.

Projects were followed in short field works over several months. Field work included interviews, meeting observations, and document studies of internal documentation on projects. Within a short period of time among data team members, I noticed that there was a lot of frustration tied to their work. Although the Norwegian public sector is one of the most digitized in Europe, and data-driven technology is pushed in a variety of policy papers, both data teams struggled to develop and implement data assemblages within their organizations. They struggled to materialize the STI of datafication, which promises that the public sector can make use of fast and cheap data analysis, glean data from entire populations, let data speak for themselves and achieve an impartial view from nowhere (Rieder & Simon 2016). The analysis demonstrates how the apparatuses and elements of the data assemblage, such as policy and strategy documents, organizational scope, legal mandates, and data and infrastructures, presented challenges and obstacles to the data teams. Letters of allocation and other internal strategy documents, for example, often decide on what is prioritized in the data teams. Specific projects must then be linked to public mandates and assessed by lawyers before data team members may even start to work. Public administration is built on legal frameworks that both enable work in the data teams and constrain what and how data practices can be enacted. Many of the observed efforts were additionally stopped at the pilot phase as the data teams did not receive enough funds to develop working systems. In addition, the members often acted as intermediaries between a variety of

actors, such as subject matter experts and technical staff. There was surprisingly little interest in data-driven technology in the public sector and the teams, therefore, spent much time presenting their work to other sections and thus actively circulating STIs. The analysis of this paper contains one of my favorite quotes from the field work: "Building the actual machine learning models is what takes the least time here (laughs)" (SKATT data team member). This again indicates a need to decenter technology in our analysis of datafication.

In this paper, I bridge the concepts of the data assemblage (as introduced by Kitchin and Lauriault 2014) and STIs to enhance our understanding of how STIs enter, are translated, and are negotiated into assemblages of materiality (Jasanoff and Kim 2015). Simultaneously, I demonstrate how STIs are always contingent, multiple, and contested, and how they often do not materialize without interruption (Mager and Katzenbach 2021; Caplan et al. 2020). The promise of a more effective and informed public administration was restricted by the technical, legal, organizational, methodological, and political constraints observed in the field work. I therefore argue that public administration seems to be a context where STIs of datafication thrive but new data assemblages struggle to materialize. Although this leads to much frustration among data team members, this also means that the constraints presented prohibit the data teams from utilizing the vastly enhanced possibilities afforded by emerging technologies to understand, predict, and control the activities of citizens. The obstacles can then serve as entry points to engage reflexively with datafication and discuss how datafication might be reconfigured. In the conclusion of this paper, I introduce the idea of the actionable critique, which I will elaborate upon in the final section of this dissertation.

5.2 Paper 2: Towards a Data-Driven Public Administration

Co-Authored with Heather Broomfield, published in the Scandinavian Journal of Public Administration.

As observed in the field work of Paper 1, the development in the Norwegian public sector was about more than introducing machine learning into public service decision-making. The second paper of the dissertation was therefore intended to provide a general overview of datafication activities within the Norwegian public sector. It aimed to demystify data-driven public administration and directed its analytical attention to critically investigate the administrative reform in the making, as called for by a variety of scholars across fields (Agarwal 2018; Brauneis and Goodman 2018; Redden 2018; Wirtz, Weyerer, and Geyer 2019). At the same time, we intended to introduce datafication to public administration studies by situating this paper within the tradition of implementation studies (Caiden 1999; DeLeon 1999; O'Toole Jr 2000) as there seems to be a severe lack of attention paid to issues of datafication within the wider field of public administration studies. Indeed, this paper is among the first to use this concept in a public administration journal, distinguishing itself from earlier research on eGovernment. Focusing on perceived challenges and unseen issues, the paper asks the overall research question of what

challenges are encountered and problematized in a nascent phase of data-driven public administration implementation. We examine what is prioritized and problematized at the policy level and among practitioners to realize data-driven ambitions and identify discrepancies in prioritizations.

The paper makes use of a multi-method approach, including a survey sent out to several public sector organizations, follow-up interviews, and document studies of policy papers, all conducted in 2019. The findings show that although data-driven is regarded as highly important across the public sector, Norwegian public administration is still in a nascent phase when it comes to realizing datafication. In addition, there is a clear lack of common understanding of what data-driven public administration is and how it should be materialized among policymakers, practitioners, and leaders in the public sector. Many organizations are in a planning, strategy, or piloting phase regarding data-driven technology as a main enabler of efficiency in the sector but remain unsure of how to make use of it. Practitioners experience a variety of challenges when attempting to become data-driven. Organizational issues, privacy and security concerns, and legal frameworks rank high among the obstacles. However, challenges also differ widely across the sector, which is especially apparent in the interviews. The analysis then goes on to show how policy papers often prioritize technical infrastructure, access, and quality of data when problematizing data-drivenness, while practitioners are more concerned with organization, internal culture, and competence issues. Both policy makers and practitioners regard insecurities tied to law, privacy, and security issues as hindering datafication work. The challenges are interwoven and mutually dependent. The paper culminates in a discussion on unseen issues.

None of the involved actors addressed the potential for the changing power dynamics between citizens and state due to the increased possibility to predict and monitor citizen behavior. Although practitioners state that they are using technology "for the best of society," there is no clear, common understanding of how this is translated into practice. The risk of unintended consequences is recognized but not problematized by practitioners, who often deal with single projects in isolation. The increased influence of private sector companies over technology production and implementation in the public sector also remains an area of silence. Although ethics is a hot topic in the field, this was not regarded as challenging by many practitioners as anonymized data is regarded as unproblematic, and many argue that ethical issues are solved through legal frameworks. Many of the classic questions of public administration dealing with equity, accountability, political legitimacy, discretion, and human agency in decision-making are not addressed by any of the actors. This reinforces the importance of public administration scholars becoming involved in the scholarly and practical debate over datafication.

5.3 Paper 3: In Search of the Citizen

Co-Authored with Heather Broomfield, published in Big Data & Society.

This paper originated in the initial analysis of the data material presented in Paper 2, where we found a surprising dearth of the terms "citizen" and "residents" in both the quantitative and qualitative data presented. This sparked our curiosity and urged us to do a secondary analysis of the data material and an investigation of the production of policy documents. Datafication has the potential to significantly change citizen-state relations. How are citizens problematized and included in policy and practitioners' discourse in the datafication of public administration? In this paper, we aimed to answer this research question by investigating who is consulted in the policy-making process and the discourse within the resulting policy documents among practitioners. Cardullo and Kitchin's (2019) scaffold of smart citizen participation served here as a blueprint for examining citizen participation in public administration datafication. We adapted this scaffold to investigate data-driven efforts on a system level. In addition, we included contextual factors in this analysis, such as the idea of Norwegian corporative pluralism, which encourages the inclusion of civil society and organized interests in policy-making and implementation through, for example, hybrid public advisory committees (Christiansen et al. 2010; Krick and Holst 2021; Rokkan 1966). This paper makes use of a multi-method approach to provide a situated analysis of datafication efforts that allowed us to research the underlying social mechanisms and the imaginaries of the agents responsible for framing and embedding the data-driven public sector.

The results of the analysis show a clear lack of citizen participation in the framing of policy papers. Although all the policy papers analyzed illustrated an extensive need for analysis and consultation with stakeholders, this is limited to consultations within the private and public sectors. Datafication was presented as both necessary and inevitable in the resulting policy papers, often linked to large societal challenges such as demographics, the oil sector's decline, and sustainability goals. Nevertheless, policy documents also placed significant emphasis on the concepts of being user-centric and needs-based. Clearly, it was mainly the private sector's needs that were considered in the policy papers. User-centered ideas are more about the solution than the process itself. Seamless services are, for example, juxtaposed with user-centricness. Here, data are imagined to flow seamlessly across organizational boundaries to provide better user experiences. Although practitioners state that their work on datafication is intended for "the good for society," this often translates into a narrow focus on privacy in projects and a vague idea of digitalization in a transparent, inclusive, and trustworthy way.

Overall, this paper found that citizens and civil society are rarely actively included in the policy-making process on data-driven public administration beyond specific and highly specialized projects. This is surprising as Norwegian corporative pluralism encourages the inclusion of stakeholders in policymaking and implementation on all levels. We argue that there is an overall top-down and paternalistic approach to datafication as these issues are often seen as both inevitable and apolitical. The context, values, and agendas embedded in data-driven public

administration are therefore obscured from citizens and civil society, and they are not actively encouraged to discuss or challenge current datafication efforts. It seems, therefore, that datafication is seen as an internal or public sector issue, not as a democratic or political one.

5.4 Paper 4: Public Sector Data as a Resource

Co-Authored with Heidrun Åm

Policy documents and government white papers are central to the datafication of public administration in Norway as they promote, frame, and enact data-driven public administration. In addition, these documents present the most definite and materialized version of the datafication STI as most projects seem to still remain pilots, as shown in Papers 1 and 2. In this final paper, we therefore studied how sociotechnical imaginaries of datafication are constructed, embedded, and maintained in the Norwegian public sector. This paper is situated within studies on technology policies and made use of Bacchi's (2009, 2012) WPR approach and Asdal and Reinertsen's (2020) practice approach to document analysis. We analyzed six documents published between 2015 and 2021, identifying the policy problems for which these technological developments seem to be the best solutions, and investigating how the STI of 'data as a resource' emerged in public administration.

Problematizations are central to modern governing processes as policy itself is actively engaged in defining and producing problems according to Bacchi (2009). By marrying the WPR approach with Jasanoff and Kim's (2015) account of STIs, we show how a variety of actors produce future visions and expectations of datafication. We began the analysis with a historical account of developments in the public sector and argued that there has a been a shift in prioritizations in recent years. The policy problem representations addressed in the documents can be divided into two major categories: the problem of public administration, where efficiency and better services are central, and the problem of data-driven innovation, where data-sharing from the public to the private sector and value creation is central. While the first policy problem addresses mostly internal public administration processes and service delivery, the second overarching policy problem addresses the whole of the economy in Norway. Both have been equally addressed in earlier policy; however, recent documents all emphasize the second problem. There is also an increased certainty in language in later policy documents.

The analysis then goes on to discuss and present actors, drivers, and justifications of datafication in the policy papers. This is about both an idea of technological drivers that force the public sector to adapt and an assumption that the sheer availability of technology such as AI and big data makes these relevant and important to public administration. The documents go on to justify action through international rankings and the implicit notion of a race between nations. Interestingly, the OECD and EU are both important actors in the policy papers analyzed. We can read such references to supranational actors as ethos-oriented argumentation, in which the truth and authority of the claims become justified through their international reach and the standing

of these institutions. These institutions also contribute to justifying datafication efforts by administrating a variety of international rankings and quantified economic analyses projecting the future revenue of data-driven public administration.

The discussion of Paper 4 once again culminates in unseen issues and issues left unproblematic. Overall, the analyzed documents all focus on enabling datafication rather than governing or constraining this development. The documents draw on a variety of success stories from the public and private sectors to legitimate this development. Regarding data as a resource that just needs to be shared and utilized renders it natural and beyond political control. While the documents address data more generally, only certain types of data are used as examples in the documents. Democratic principles and societal issues are not addressed in any of the analyzed documents as regulation and welfare distribution efforts are left unseen. This is a surprising finding as the Norwegian welfare state has long traditions of collectively distributing benefits from common resources (Ryggvik 2009; Thue 2003).

6 Discussion and conclusion

Although each of the four papers presents an individual contribution to the scholarly debate on the data-driven public sector, as elaborated upon in the previous chapter, an additional contribution of this dissertation lies in the intersection and synthesis of all four papers. I set out to investigate how datafication is framed and embedded in Norwegian public administration by studying the inner workings of the data-driven public sector. I begin this concluding section by highlighting the importance of a situated practice approach to datafication. I wish to emphasize challenges and silences as examples of how a situated practice approach can help us identify new avenues of citizen intervention or contestation. The discussion will culminate in reflections on how critical data and algorithm studies might work toward an actionable critique of public administration datafication.

6.1 Situating datafication in public administration

In the introduction section of this dissertation, I presented several shortcomings regarding the current state of research on datafication in the social sciences. There has been an overwhelming focus on the private sector application of data-driven technology and its consequences for society. This body of literature often treats datafication as something that is already being used in all aspects of social life. As argued in Papers 1, 2, and 4, the data-driven public sector is still in a nascent phase of development and remains, in many cases, a future vision. These visions and imaginaries are, however, still performative and deserve scholarly attention. Although an emerging body of literature on the public sector application of data-driven systems is evident, empirical work is often confined to case studies on single projects within specific sectors or organizations in the public sector. This dissertation aimed to research datafication at different levels to understand how it is embedded and framed in the Norwegian public sector and to ask what practitioners, institutions, and policy makers do (not) and say in relation to datafication. Asking these questions allowed me to assert the political dimension of datafication and "point to both the (re)construction of citizens in data systems and the emerging opportunities for citizen intervention and resistance" (Dencik 2019:244).

One of the key ambitions of critical data and algorithm studies is to illuminate the various processes in which our lives are turned into data and then recirculated into action, so as to question the assumptions they are built upon and critically intervene in their deployment (Beer 2016; Iliadis and Russo 2016; Kitchin and Lauriault 2014; Ruppert et al. 2017). Through this dissertation, I wish to demonstrate that critical data and algorithm studies benefit greatly from a situated practice approach to datafication, as introduced by Dencik (2019). Following Bakardjieva (2020), I furthermore want to argue that practice theory is best used here as a sensitizing concept—a reference and guidance for approaching empirical instances—that can be supplemented with a variety of other concepts and approaches to enhance our understanding of

the data-driven public sector. A situated account of datafication shifts the focus of data or data harms away from entry points of contestations and toward the importance of understanding the interests, agendas, and power relations in which datafication is embedded. Rather than focusing solely on the immediate consequences and impact of datafication, the situated practice approach places this phenomenon within a wider landscape of practices. Paper 1 is based on fieldwork among data teams in two Norwegian public sector organizations and offers valuable insights into how data team members struggle to align new data practices with existing practices and material infrastructures in their organizations. Paper 2 then highlights the tensions that arise when datafication enters public administration more generally, as well as the challenges experienced by practitioners. In Paper 3, we included two key social and democratic practices in our analysis of how citizens are problematized and included in datafication, the corporate pluralism model that has a long tradition in Norwegian policy making (Rokkan 1966), and hybrid advisory committees (Krick and Holst 2021). We argued that both practices are disregarded in the constructing and framing of datafication, as datafication is presented as apolitical and, therefore, seemingly not in need of the inclusion of citizens and civil society in policymaking; citizens are not able to challenge or mediate datafication through informal or formal practices of participation in public administration. Paper 4 then studies how the STI of data as a resource has emerged and is embedded into the Norwegian public sector and what problematizations this is built upon. In this paper, we argue that although the Norwegian public sector has previously focused its practices on collectively distributing the benefits of resources (Ryggvik 2009; Thue 2003), democratic arguments are solely used to make public data available for business development. The relationship between existing public administration practices and datafication has thus been explored in several of the papers.

As argued by Bareis and Katzenbach (2021), although the STIs of datafication often resemble each other, they also exhibit distinct local features. This points to the importance of context in producing a critique of datafication. Again, much of the empirical and conceptual work on issues of datafication is still done in the context of private internet platforms and social media. The critique of datafication produced in this context is of great importance to both academic and societal discussions, but as demonstrated in all four papers, datafication needs to be situated beyond internet platforms. We need to challenge universal assumptions about datafication and look more closely at how it is translated into specific contexts (Milan and Treré 2019; Wijermars and Makhortykh 2022). Public administration is different; it operates under the bottom line of creating public value—however this might be interpreted—and is highly regulated by law. The public sector already operates with thousands of data assemblages and has a long history of data collection, management, and use (Desrosières 1998; Porter 1996). Furthermore, Norwegian public administration differs from other public administration organization models. The Norwegian public sector keeps vast registers on its citizens, who cannot opt out of these practices of data collection and management. While Norway is one of the most digitized countries worldwide, the Norwegian public sector also enjoys high levels of trust among its citizens. As

Mergel et al. (2019) pointed out, how digital transformation and other administrative reforms are perceived and interpreted is strongly influenced by different bureaucratic traditions. Even within public administration, we can identify significant variations regarding how and why datafication is approached. Paper 1, for example, shows how the SKATT data team struggled less with producing machine learning than the NAV data team because of its relatively wide control mandates and long tradition of data analysis and if/then automation. Paper 2 demonstrates how practitioners across the Norwegian public sector interpret datafication differently and how each organization has a distinct and unique set of perceived challenges that is tied to their organization, operation, and existing data practices. There is, therefore, a need to produce context-specific accounts of how datafication is understood "as well as to question the assumptions on which its evaluation within different domains is based" (Wijermars and Makhortykh 2022:947). This is about understanding and taking into account the distinct apparatuses of public administration data assemblages when researching the data-driven public sector.

A situated practice approach to datafication decenters technology and researches the values and interests that pertain to data in different contexts. It pays special attention to the institutional practices and material resources of specific environments that shape the new social order under datafication: in public administration. In this dissertation, I have operationalized this approach by paying analytical attention to the apparatuses of the data assemblage, the STIs of practitioners and policy makers, and how these are constructed and embedded into public administration practice. The concepts of the data assemblage and STIs have here been used both as methodological lenses and analytical tools and have assisted in implementing a practice approach, as introduced by Couldry (2004, 2012) and Dencik (2019, 2020). This was achieved by addressing the following research questions: What interests and agendas are inscribed into the STI of datafication, and what are omitted? How has this imaginary been constructed and embedded into the Norwegian public sector? What obstacles do practitioners experience when attempting to stabilize STIs into working data assemblages?

Operating within a new and emerging subfield within the social sciences comes with several challenges. As there were only a few studies available on public administration datafication in 2018 and 2019, I initially struggled to pinpoint a specific research gap because there seemed to be more gaps than was possible to grasp. This required an explorative research design. I benefited greatly from abductive analysis and its emphasis on elements of surprise (Tavory and Timmermans 2014; Timmermans and Tavory 2022). I used elements such as the observation that AI just seemed to be one of many developments in the public sector circling around data and the absence of successful projects in the observed organizations strategically to steer and confine my work. A situated approach allows us to ask new questions and follow our subjective yet theoretically informed moments of surprise. The sub-questions developed throughout the project are thus just one of many ways to approach datafication.

In working on finding an appropriate scope for this dissertation, I have developed an understanding of public administration datafication, as follows:

> Data-driven administration consists of two interwoven processes: the use of more and/or different kinds of data and the recirculation of data through ever more complex methods, such as machine learning or automatic decision-making.

This definition is central to all papers and attempts to promote an understanding of datafication that goes beyond simple notions of quantification and extractivism. However, as argued across papers, how datafication is understood and translated into practice varies across time and space. In Paper 4, we demonstrated, for example, that while improving services through internal data sharing was highlighted in policy papers analyzed earlier, we observed an increasing interest in the commodification of public data in later policy papers. Therefore, I suggest that we need to move beyond a static and linear understanding of the data-driven public sector and be sensitive to the contested and multiple nature of datafication processes across time and space.

Although the situated practice approach does not focus on the immediate consequences of datafication, it does not render the data-driven public sector harmless. As Dencik (2020:251) stresses, "the extent to which citizens are able to challenge, avoid or mediate their data doubles—that is the relation between data practices and other social practices—becomes a key political question of our time." In this final discussion, I want to briefly reflect on the opportunities of citizen intervention, contestation, and resistance, as these emerge across papers. A situated practice approach actively researches and discusses potential entry points for contestations to avoid determinism. In this dissertation, I have explored two such entry points: **obstacles and silences.**

6.2 Challenges and obstacles

What do practitioners, institutions, and policy makers do in relation to datafication? First and foremost, they seem to struggle, as is especially apparent in Papers 1 and 2. My empirical work has repeatedly encountered frustration among practitioners, which led to one of many moments of surprise: this seemed to correspond little with previous empirical work on datafication on internet platforms, where the myth and hype about datafication often seemed to obscure the many struggles of making data-driven technology work (Elish and boyd 2018; Hagendorff and Wezel 2020). In addition, the work observed seemed to misalign with the imaginaries identified in previous research, where the public sector can make use of fast and cheap data analysis, glean data from entire populations, let data speak for themselves and achieve a impartial view from nowhere under datafication (Rieder & Simon 2016). Many researchers seemed to assume that the STIs of datafication materialize without interruption across society (Caplan et al. 2020). The findings presented in Papers 1 and 2, however, clearly demonstrate that this is not the case. Following Dencik's (2019) suggestion that we need to pay close attention to the multiple, converging, and conflicting forces that shape the continuously constructed process of

datafication, I have consciously focused on tensions in this dissertation. This was also part of my ambition to present alternative narratives about datafication and move beyond the hype of AI and big data. While researchers such as Zejnilović et al. (2020) have focused on the frictions between data-driven systems and street-level bureaucrats using these systems in casework and welfare provision, I have concentrated on the tension that can be observed when data-driven technology and the idea of the data-driven public sector are embedded and produced in public administration prior to systems being used in actual casework or in internal processes.

While entering public administration through my fieldwork at NAV in 2018 and 2019, I struggled to find working systems beyond simple and confined pilot projects. Although conferences and policy papers all stressed the importance of becoming data-driven, little seemed to materialize in practice. This continued throughout my fieldwork and was also a recurring topic in the AI fagforum. I observed small and confined projects or pilots, but only a few systems that were doing work automatically every day. Rather than only regarding the apparatuses of the data assemblage as a set of contingent, relational, contextual, discursive, and material relations, I argue in Paper 1 that we might regard the elements and apparatuses of the data assemblage as mediators and possible counter forces. Focusing on obstacles allowed me to show the distinct features of public administration data assemblages. I show how the two data teams researched face different challenges based on, for example, existing material infrastructures and data practices, organization-specific laws, and a focus on differing subject matter expertise.

Paper 2 follows this focus on obstacles and challenges and presents a general mapping of the Norwegian public sector in which we introduce a variety of perceived challenges that public sector practitioners encounter when trying to embed datafication into their everyday work practices. This includes issues such as legal assessments, data quality, data access, and a lack of financing. We also find that many of the challenges are drawn up in policy documents. However, what is problematized differs greatly between policy makers and practitioners. While policy papers seem to simply regard datafication as a matter of making data and infrastructure available and to give more leeway in the law to unlock the effectivization potential, practitioners struggle with a variety of obstacles beyond these issues.

The focus on challenges has several aims in the work of this dissertation. First, the outlined challenges clearly demonstrate that datafication rarely materializes fully and without interruption in public administration. As argued in Paper 1, public administration seems to be a context in which the STIs of datafication thrive but struggle to materialize. This once again reinforces the importance of situating datafication in public administration, paying attention to its distinct practices and apparatuses. By focusing on challenges and obstacles, we draw attention toward the often-invisible work involved in assembling data assemblages. This also redirects our analytical focus away from researching the negative outcomes of datafication and toward the tensions between existing practices and values and the datafication paradigm.

Second, the challenges of technology production and implementation are popular topics within public administration scholarship (Agarwal 2018; Chen and Hsieh 2014). As demonstrated in Paper 2, many of the challenges experienced by practitioners resemble earlier research on administrative reform (Kraemer and King 2006; O'Toole Jr. 2000). Challenges can therefore serve as boundary objects between critical data and algorithm studies and public administration scholarship to facilitate dialogue between the fields, as demonstrated by the publication of Paper 2 in a public administration journal. Third, challenges are of great interest to practitioners in the field. Both Papers 1 and 2 show how practitioners experience much frustration in their everyday work on datafication, as there are a variety of obstacles hindering work. I have used challenges in several of the presentations to practitioners to find common ground from which to discuss datafication more generally. The visualization and focus on the complexity of interlinked challenges has led to interesting discussions with practitioners.

Finally, as argued in Paper 1, these challenges can also be seen as contestations and aid us to reflexively engage with public administration datafication. Why do projects fail? An easy explanation often cited in the interviews contained in Papers 2 and 3 was the internal resistance to change within public administration or the lack of data readily available for analysis. This reflects the public administration literature, which often focuses on the uncertainties and readiness of the sector to adopt new technology (Guenduez et al. 2020; Kim, Trimi, and Chung 2014; Klievink et al. 2017). This literature attempted to explain failures or non-adoptions by, for example, focusing on the lack of common understandings among public managers (Guenduez et al. 2020). A more complex explanation might be that data-driven systems have the potential to fundamentally change public administration and social order. As Hintz et al. (2018) argued, datafication enhances the state's ability to predict, control, and classify citizen activities, thereby changing power dynamics. Norwegian public administration has installed a variety of both formal and informal mechanisms to safeguard some of the core principles of the welfare state, which I argue often become visible in the challenges encountered by practitioners. Engaging reflexively with obstacles and challenges might also help us understand how we can counter the datafication of public administration and produce actionable critiques of the phenomenon. It is concerning that both policy makers and practitioners actively engage in identifying and eradicating obstacles in the public sector. As all papers show, there is a sense of urgency to enable datafication within public administration, which is often translated into eliminating all challenges and obstacles rather than critically engaging with them. This leads us to the issue of silences in policy and practice.

6.3 Silences

As argued in the theory chapter of this dissertation, I regard it as highly important to research what people, institutions, and policy makers do not do and say in relation to datafication, as these silences and omitted issues might point us toward asking new questions and broadening the discourse on datafication in public administration. Focusing on what is left unproblematic or what

and who is silenced helps me explore the critical potential of the situated practice approach. As boyd and Elish (2018) argue, the discourse of myth and hype related to datafication often disguises problematic issues of data and technology and therefore contributes to the disguising of the power of data-driven systems. My work on answering the overall research question of how datafication is framed and embedded in the Norwegian public sector has uncovered a clear misalignment between the practicing field and critical data and algorithm studies. While the practicing field and policy documents seem to closely resemble public administration scholarship's focus on enabling datafication, critical data and algorithm studies have pointed out a variety of problematic issues of datafication that often do not appear in practitioner and policy discourse.

Focusing on silences was initially inspired by Redden's (2018) counter-mapping of the Canadian public sector, where she identified two significant areas of silence: the impact of public-private partnerships in datafication efforts and a lack of addressing how data-driven systems may change the way citizens are governed and understood. In Paper 2, we take this as a point of departure and identify several omitted issues, including the disregarding of changing citizen-state relations and a lack of oversight within the public sector on what is done and where. Practitioners regard ethics as something that can be solved through legal frameworks. There was little consideration of the unintended social consequences of datafication and the impact of automation on the role of discretional decision-making in public administration, which was surprising, as there are several public examples of harmful systems in Europe, such as the Swedish Trelleborgmodellen or the famous Dutch fraud detection model. We argue that many of these omitted issues are far more complex than technical or legal issues drawn up in policy documents; therefore, they often receive little attention in practice and policy. This is then also paired with an overall paternalism in public administration, under which practitioners and policy makers argue that their work is always for the best of society, without having a concrete understanding or idea of what this may contain. Citizen perspectives are one specific issue that remains unseen by the practitioners and policymakers identified in Paper 2. In Paper 3, we take this silence as a point of departure to further investigate how citizens are involved in making policy, and thus how they are included in constructing and embedding the STI of datafication. What we find is a paternalistic and top-down approach to datafication, under which the presence of citizens seems to be regarded as unnecessary in policy-making processes. Although the Norwegian public sector encourages user involvement in projects, this is mostly limited to the late stages of projects and very confined areas. We argue that this is because datafication is regarded as inevitable, apolitical, and problem-free by many practitioners and policy makers, and we try to challenge this understanding in the paper, as citizens and civil society seem to be actively silenced in the pursuit of public administration datafication.

Paper 4 continues to argue that policy papers often consider data as neutral and beyond political control. These documents often argue that what can be controlled are the varieties of obstacles

currently hindering datafication. Interestingly, the policy papers focused on enabling datafication rather than regulating or governing datafication and once again fail to address the potential negative societal impacts of data so often drawn up in critical data and algorithm studies. Although the Norwegian public sector has established practices to safeguard the collective distribution of benefits from common resources (such as public administration data), this is often disregarded in the datafication discourse. The public sector's responsibility confined to make data available and eradicate any obstacles that hinder datafication today, to avoid 'lacking behind'. The inherent uncertainty tied to the STI of data as a resource is masked by calculations and international comparisons, again establishing a deterministic view of datafication and an idea of inevitability. This renders alternative ways of imagining the future absent from the analyzed documents. Although STIs are often multiple and contested (Mager and Katzenbach 2021; Ruppert 2018), contestations become difficult when facing this complex network of justifications, coupled with increased discursive certainty. The documents establish authority through constantly referring to themselves and drawing on a variety of actors to do so. Again, going back to Paper 3, we can also see that several societal actors are fully excluded in producing these policy papers and possible contestations or alternative ways of imagining the future might therefore be omitted or silenced in documents.

An unseen issue I find particularly interesting is the absence of any mentions of (partly) failed attempts at public administration datafication. Indeed, one of the motivations for writing Paper 4 was to understand why public administration continues to encourage datafication through publishing policy papers when various projects seem to meet significant obstacles. As Paper 4 shows, there is an overwhelming focus on success stories of datafication in policy documents, often drawing on real-life yet isolated example applications of data-driven systems (in early policy papers these examples were often 'imported' from very different contexts). Initially, I aimed to follow a specific project on data-driven assessment of sickness absence from its origins through implementation and then observe it working in the world to be able to follow Dencik's (2019) suggestion to research what users and citizens do in relation to datafication. However, I had to abandon this idea, as the project was delayed repeatedly. It simply did not fit into the narrow timeframe of a PhD. However, this is not the only example of delayed or restricted projects. Difficulties in producing data-driven technology in public administration often gain little attention from either policy makers or practitioners, as the STI of public administration datafication depends on success stories to function. Therefore, it also becomes important for researchers to illuminate this silence by showing the various challenges and failed attempts at datafication and reflexively engaging with what these might tell us about the relation between existing and emerging (data/social) practices.

Focusing on silences, I had a clear ambition on broadening the discourse on public administration datafication and present alternative problematizations, without simply pointing to negative outcomes of data-driven technology (such as bias and discrimination) in isolation. As Schiølin

(2020, 545) argues, powerful STIs, such as datafication, "risk curtailing the possibility of imagining alternative futures, impede public reasoning about how to inhabit them, and favor some modes of being over others". By drawing attention to alternative ways of imagining the future through silences and omitted issues in policy papers and practitioner's discourse, I was able to critically investigate current datafication efforts, again focusing on the political dimension of this administrative reform. The outlined silences seem to point toward a misalignment between critical data and algorithm studies and public administration practitioners in what datafication is and how it is discussed. This then leads us to the last point I would like to make in this dissertation: How can we as social scientists work toward an actionable critique of datafication?

6.4 Toward an actionable critique of public administration datafication

> In academic circles there is often a discussion about the supposed tradeoff between rigor and relevance. Some fear that more emphasis on the societal value of research could interfere with academic robustness. [...] We believe that the two are intertwined and reflecting on the societal value of our research can help us to strengthen the robustness of our work by enhancing the quality of the ways in which we communicate our research objectives, methodology and findings. (Meijer and Webster 2022:1)

Following this argument, I want to use the last pages of the introduction to public administration datafication to reflect on the societal value of my work. In the theory chapter of this dissertation, I introduced four puzzles that social scientists face when researching public administration datafication: avoiding determinism in the analysis of data-driven public administration, balancing between discourse/hype and materiality/real world in studying public administration datafication, balancing between critical accounts of specific applications of data-driven technology in the public sector and a general critique of the datafication of public administration, and balancing between reactive and proactive critiques of datafication. I position my research at the intersection of these puzzles, approaching datafication as a continuous process aiming to alter existing data practices and introduce new ones to public administration. I focus on both the framing and embedding of datafication as a multitude of practices that can be observed and researched by social scientists through a multi-method explorative research design. When positioning my research based on the four presented puzzles, I had the overall ambition to make the findings of the dissertation relevant to the practicing field of public administration—in other words, to produce an actionable critique of datafication. This normative ambition has always influenced what I have done directly or indirectly, and therefore deserves to be elaborated upon in this concluding section of the dissertation. It follows Andrews (2019) suggestion to avoid a gotcha-style of research and work in close collaboration with the practicing field.

The Oxford English Dictionary defines "actionable" as "able to be acted upon or put into practice; useful, practical" (OED Online 2022). It is the practical value of the critique produced through this dissertation, in addition to its potential for being acted upon, that is central to my work—both

influenced the choice of research design, theoretical framework, and questions asked. As Bastow, Dunleavy, and Tinkler (2014:272) pointed out, the social sciences predominantly reconfigure by

> [...] influencing people to think about things in a different, more precisely reasoned and better-informed way, one that will (hopefully) produce better decisions and societal outcomes than would have been achieved without the presence of these disciplines.

This ambition to influence public administration to think differently about datafication and produce better or different decisions when it comes to the framing and embedding of datafication is at the heart of the actionable critique. This is based on the understanding that theoretical and practical work do not rule out each other. What, then, is an actionable critique of public administration datafication? An actionable critique of datafication is social constructivist in its epistemology. As Seaver (2019:413) pointed out, "The point of declaring something a construction is to argue that it might be constructed differently." This means that the critique questions the determinism often present in the field and offers contestations.

The actionable critique is situated in and focuses on the practices and imaginaries of people working through, on, and with datafication. Furthermore, the situatedness and research focus on contexts are of high importance. Although my own work is inspired by and builds on researchers focusing on private sector applications of datafication (for example, Beer 2016; Bucher 2012, 2017; Gillespie 2014b), I argue that we cannot produce an actionable critique of datafication without paying attention to the context within which it is embedded. It is in the distinct apparatuses of the public sector data assemblage that new contestations can be found. If we apply the same critique to public administration datafication as to the datafication of internet platforms, we falsely assume that datafication enfolds and materializes fully in all aspects of society and, therefore, run the risk of overdetermining its power and buying into the myth and hype surrounding big data and AI. This is in addition important, as the public sector increasingly seems to distance itself from the private sector origins of datafication, as argued in Paper 4. Several of the papers show that datafication is often in tension with public sector apparatuses and existing practices. Ultimately, we need to avoid a substantivist idea of datafication embodied with the values of rationalization and domination over society, as expressed by members of the Frankfurt School. Following this, I also want to argue against an understanding of datafication as simply yet an iteration of neoliberalism. We need to actively argue against the determinist understanding of practitioners and policy that I have pointed out in Papers 3 and 4, where datafication is seen as the inevitable future of public administration and provide alternative narratives.

In addition to context, I argue that temporality is an important aspect of actionable critique. Data activism studies often focus on mitigating harm from existing systems (Lehtiniemi and Ruckenstein 2018; Milan 2017). Contrary to this, I consciously researched datafication in the making and the practices around constructing and embedding STIs and data assemblages. Much research on data activism and resistance to algorithms directs its analytical attention toward

reactive measures or user agency (Milan 2017; Meng and DiSalvo 2018; Velkova and Kaun 2021). This body of work offers valuable insights into how citizens and users react to and resist the harmful effects of datafication. I want to highlight, however, that we also need to pay close attention to the agency of citizens in framing, constructing, and embedding datafication into society proactively. This requires us, as researchers, to engage with the STIs of datafication and the messy world of producing data-driven technology. It requires us to engage in the area of tension between discourse/hype and real world/materiality, not always sure what we are actually dealing with. In many ways, I have taken a step back with this dissertation, often refraining from discussing the immediate consequences of public administration datafication or data harm and focusing on what people do in relation to datafication in its nascent phase. This is especially important, as many of the datafication visions remain future visions. They are not yet materialized, but they still have performativity. As we conclude in Paper 2 "given that this is still in the nascent phase, there is time to adjust the course."

A situated critique of datafication should enable researchers and citizens to find emerging opportunities for intervention and resistance (Dencik 2019), rather than simply focusing on pointing out data threats. As Velkova and Kaun (2021:523) stressed, "Most research has focused on negative outcomes, including ethical problems of machine bias and accountability, little has been said about the possibilities of users to resist algorithmic power." These might be found in the unseen issues outlined in Papers 2 and 4. By communicating and making these silences visible, these papers introduce new ways of approaching and discussing datafication to policy makers and practitioners. Papers 1 and 2 outlined the various obstacles of datafication that might be used to critically and reflexively engage with datafication, thus providing an alternative to the dominant narratives of AI and big data hype. Paper 3 argues that there is a lack of citizen participation in policy development and implementation in Norway and an all-encompassing paternalistic approach to the problem and attempts to problematize this issue. What if we started including citizens and civil society in the framing of data-driven public administration? Furthermore, Paper 4 deconstructed the STIs of datafication and showed how policy has a narrow focus on efficiencies and digital markets in its work toward a data-driven public sector. How might we introduce alternative imaginaries of data and technology into society? By deconstructing and critically investigating the justifications used to promote and embed datafication, we might produce a more precise critique of the phenomenon and question the taken-for-granted assumptions of these documents.

Producing an actionable critique comes with the responsibility of feeding findings back into the public sector. As Moats and Seaver (2019) point out, social scientists often struggle to bring important insights into the practice of data science. Drawing on an editorial by Cathy O'Neil (2017), Moats and Seaver argue that critical data and algorithm studies need to become more interesting or useful to data scientists. However, as argued in this dissertation, data scientists are just one of many groups of actors involved in the process of public administration datafication. I therefore

also wish to stress the importance of critical data and algorithm studies becoming more interesting and useful to public sector practitioners of all kinds, policy makers, and citizens. Critical data and algorithm studies should emphasize an applied and practical approach according to Iliadis and Russo (2016), identifying social data problems, contributing to enhancing data literacy and data justice, and providing citizens with the knowledge to question data practices. Public administration research has a long tradition of working in close collaboration with its research object (Yang and Miller 2008; Haverland and Yanow 2012). Taking into account both fields, making knowledge available to practitioners, and trying to understand their contexts have been of great importance to my research.

Inspired by the turn toward critical collaborative ethnography (Lassiter 2005; Lassiter and Eric 2005), in which research subjects are included not only in the data collection but also in the whole of the research project, I have actively fed back research findings into the practicing field of public administration, discussing my findings and observations with both policy makers and data teams in meetings and presenting my work at a variety of conferences (an example presentation can be found in Appendix C) Like Moats and Seaver (2019), I have also found that many practitioners are far more critical of their tools and the limitations of datafication than is often assumed in critical literature. In my first presentations, I simply presented research from critical data and algorithm studies. My arguments were often dismissed with a side note of "but we are not Facebook or the Chinese social credit system." The practitioners did not see any relevance in my accounts. However, when I started to present findings from my own research on the Norwegian public sector and connected these to more general critiques of data and technology in society, my research was often received with much discussion and dialogue with and among practitioners. In these presentations, I focused on establishing a common understanding of data-driven public administration with the audience and drew up practical challenges (as presented in Papers 1 and 2) and unseen issues (as presented in Papers 2, 3, and 4) before highlighting the importance of understanding that producing data-driven technology in the private sector is not the same as producing data-driven technology in the public sector. I was invited to ever more events (and will continue to attend those), often explicitly asked to question their perception of datafication. In many ways, this has challenged and proven the "actionability" of my critique. The presentations had the clear normative ambition of broadening the discussion on public administration datafication and putting Bastow, Dunleavy, and Tinkler's (2014) proposition into action. Although many of the findings presented in these presentations might seem mundane or even trivial to scholars within critical data and algorithm studies, they were not perceived as such by the intended audience. I think there is a great but often-unexplored potential for social scientists to critically investigate the datafication of public administration and then to engage with both the practicing and scholarly fields of public administration.

6.5 Avenues for further research

This research project presents a partial and situated account of how datafication is framed and embedded into Norwegian public administration. It presents a first empirical and explorative encounter with the public sector and offers various avenues for further research. As pointed out, we need to recognize that public administration differs greatly from the private sector and across countries to produce more precise critiques of datafication, yet also recognize the similarities of datafication practices across contexts. There is a need for comparative case studies between countries to identify the unique local features of what people and institutions do and say in relation to datafication and to enhance our understanding of how material infrastructures and existing practices shape data-driven efforts. In other words, identifying similarities and differences between the apparatuses of data assemblages. Comparative research should then also include studies beyond the American and European standpoint, as advocated by researchers such as Liu (2022) and Milan and Treré (2019).

As datafication efforts in the Norwegian public sector progress, it is important to continue empirical work on the inner-working of the data-driven public sector to investigate how datafication enters and is embedded into the day-to-day work practices of case workers, street-level bureaucrats, and other operations of the public sector. Here studies on emerging and changing citizen state practices are of great importance. Following the work done by researcher such as Kennedy et al. (2020) and Hintz et. al (2022) these empirical encounters should then also pay close attention to the multiple ways in which citizens might react, feel, alter, and resist the data-driven public sector.

I started this introduction to public administration datafication with a quote about how the massive success of Facebook and other internet platform companies have inspired the public sector in Norway to embed the STI of datafication. However, it seems that the public sector has since detached itself both discursively and practically from datafication's' origins in the private sector. We therefore need more research on the specific ways in which the public sector has translated and taken ownership over the idea of data as a resource, as well as the justifications and practices associated with the datadriven public sector. As I have here focused on national efforts of datafication, there is a need to deepen our critical understanding of how these local efforts are tied to international advocates such as the European Union and the OECD.

The datafication of public administration remains an under-researched and emerging object of inquiry for the social science. Indeed, at the end of conducting this research project, I am left with more questions than ever. As Kitchin (2014:186) points out, we need "deep, careful, and critical reflection [...] putting theory to work through empirical case studies". Through this dissertation I have hopefully demonstrated the value of situated practice approaches to datafication that can inspire other social scientists to engage with the messy inner workings of public administration datafication. What is required then are deep, careful, and critical engagements with what citizens, practitioners and policymakers imagine, say, and do in relation to datafication across both the

public and private sector. More interdisciplinary work is needed to help us recognize the social and political dimension of public administration datafication and find new ways of citizen resistance and contestation.

7 Literature

Agarwal, P. K. 2018. "Public Administration Challenges in the World of AI and Bots." *Public Adm*inistration Review 78(6):917–21.

Agger, Ben. 1991. "Critical theory, poststructuralism, postmodernism: Their sociological relevance." *Annual review of sociology* 17:105-131.

Alcaide–Muñoz, Laura, Manuel Pedro Rodríguez–Bolívar, Manuel Jesús Cobo, and Enrique Herrera–Viedma. 2017. "Analysing the Scientific Evolution of E-Government Using a Science Mapping Approach." *Government Information Quarterly* 34(3):545–55.

Allhutter, Doris, Florian Cech, Fabian Fischer, Gabriel Grill, and Astrid Mager. 2020. "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective." *Frontiers in Big Data* 3(5):1-17.

Ananny, Mike, and Kate Crawford. 2018. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20(3):973–89.

Andreassen, Rikke, Anne Kaun, and Kaarina Nikunen. 2021. "Fostering the Data Welfare State: A Nordic Perspective on Datafication." *Nordicom Review* 42(2):207–23.

Andrejevic, Mark, Lina Dencik, and Emiliano Treré. 2020. "From Pre-Emption to Slowness: Assessing the Contrasting Temporalities of Data-Driven Predictive Policing." *New Media & Society* 22(9):1528–44.

Andrews, Leighton. 2019. "Public Administration, Public Leadership and the Construction of Public Value in the Age of the Algorithm and 'Big Data.'" *Public Administration* 97(2):296–310.

Appadurai, Arjun. 1996. Modernity at large: Cultural dimensions of globalization. Minneapolis: University of Minnesota Press.

Arnoldi, Jakob. 2016. "Computer Algorithms, Market Manipulation and the Institutionalization of High Frequency Trading." Theory, Culture & Society 33(1):29–52.

Arter, David. 2004. "Parliamentary Democracy in Scandinavia." *Parliamentary Affairs* 57(3):581–600.

Asdal, Kristin, and Hilde Reinertsen. 2020. *Hvordan Gjøre Dokumentanalyse: En Praksisorientert Met*ode. Oslo: Cappelen Damm Akademisk.

Atkinson, Paul, and Lisa Morriss. 2017. "On Ethnographic Knowledge." Qualitative Inquiry 23(5):323–31.

Baack, Stefan. 2015. "Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism." Big Data & Society 2(2):1-11.

Bacchi, Carol. 2009. *Analysing Policy.* Frenchs Forest: Pearson Higher Education AU.

Bacchi, Carol. 2012. "Why Study Problematizations? Making Politics Visible." *Open Journal of Political Science* 2(1):1-8.

Bakardjieva, Maria. 2020. "Practicing Media—Mediating Practice| New Paradigm or Sensitizing Concept: Finding the Proper Place of Practice Theory in Media Studies." *International Journal of Communication* 14:2928-2945.

Barassi, Veronica. 2019. "Datafied Citizens in the Age of Coerced Digital Participation." *Sociological Research Online* 24(3):414–29.

Bareis, Jascha, and Christian Katzenbach. 2021. "Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics." *Science, Technology, & Human Values* 47(5):855-881.

Bastow, Simon, Patrick Dunleavy, and Jane Tinkler. 2014. *The Impact of the Social Sciences: How Academics and Their Research Make a Difference*. London: Sage Publications Ltd.

Beer, David. 2016. "How Should We Do the History of Big Data?" *Big Data & Society* 3(1):1-10.

Beer, David. 2017. "The Social Power of Algorithms." *Information, Communication & Society* 20(1):1–13.

Beira, Eduardo and Andrew Feenberg. 2018. Technology, Modernity, and Democracy: Essays by Andrew Feenberg (Reinventing Critical Theory). London: Rowman & Litfield Publishers

Benbouzid, Bilel. 2019. "To Predict and to Manage. Predictive Policing in the United States." *Big Data & Society* 6(1):1-13.

Bijker, Wiebe E., and John Law. 1994. *Shaping Technology/Building Society: Studies in Sociotechnical Change*. London: MIT press.

Blumer, Herbert. 1954. What is wrong with social theory. *American Sociological Review* 19(1):3-10.

Bowker, Geoffrey C. and Susan Leigh Star. 1999. *Sorting Things Out: Classification and its consequences*. Cambridge, MA: MIT Press.

boyd, Danah, and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15(5):662–79.

Brauneis, Robert, and Ellen P. Goodman. 2018. "Algorithmic Transparency for the Smart City." *Yale JL & Tech.* 20:103-176.

Bucher, Taina. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media & Society* 14(7):1164–80.

Bucher, Taina. 2017. "The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms." *Information, Communication & Society* 20(1):30–44.

Bullock, Justin B. 2019. "Artificial Intelligence, Discretion, and Bureaucracy." *The American Review of Public Administration* 49(7):751–61.

Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3(1):1-12.

Caiden, Gerald E. 1999. "Administrative Reform--Proceed with Caution." *International Journal of Public Administration* 22(6):815–32.

Callon, Michel. 1984. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay." *The Sociological Review* 32(1):196–233.

Campbell-Verduyn, Malcolm, Marcel Goguen, and Tony Porter. 2017. "Big Data and Algorithmic Governance: The Case of Financial Practices." *New Political Economy* 22(2):219–36.

Caplan, Robyn, and danah boyd. 2018. "Isomorphism through Algorithms: Institutional Dependencies in the Case of Facebook." *Big Data & Society* 5(1):1-12.

Caplan, Robyn, and Martin Partin. 2020. Against Platform Determinism: A critical orientation.

Cardullo, Paolo, and Rob Kitchin. 2019. "Being a 'Citizen' in the Smart City: Up and down the Scaffold of Smart Citizen Participation in Dublin, Ireland." *GeoJournal* 84(1):1–13.

Cath, C., S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi. 2018. "Artificial Intelligence and the 'Good Society': The US, EU, and UK Approach." Science and Engineering Ethics 24(2):505–28.

Cavanillas, José María, Edward Curry, and Wolfgang Wahlster. 2016. *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*. London: Springer Nature.

Chaudhuri, Bidisha. 2022. "Programmed Welfare: An Ethnographic Account of Algorithmic Practices in the Public Distribution System in India." *New Media & Society* 24(4):887–902.

Charmaz, Kathy. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. London: Sage.

Chen, Yu-Che, and Tsui-Chuan Hsieh. 2014. "Big Data for Digital Government: Opportunities, Challenges, and Strategies." *International Journal of Public Administration in the Digital Age* 1(1):1–14.

Cheney-Lippold, John. 2017. "We Are Data." in *We Are Data*. New York: New York University Press.

Choroszewicz, Marta, and Beata Mäihäniemi. 2020. "Developing a Digital Welfare State: Data Protection and the Use of Automated Decision-Making in the Public Sector across Six EU Countries." *Global Perspectives* 1(1):12910.

Christensen, Tom, Amund Lie, and Per Lægreid. 2008. "Beyond New Public Management: Agencification and Regulatory Reform in Norway." *Financial Accountability & Management* 24(1):15–30.

Christiansen, Peter Munk, Asbjørn Sonne Nørgaard, Hilmar Rommetvedt, Torsten Svensson, Gunnar Thesen, and PerOla Öberg. 2010. "Varieties of Democracy: Interest Groups and Corporatist Committees in Scandinavian Policy Making." *Voluntas: International Journal of Voluntary and Nonprofit Organizations* 21(1):22–40.

Christin, Angèle. 2017. Algorithms in practice: Comparing web journalism and criminal justice. Big Data & Society 4(2):1-14.

Cieslik, Katarzyna, and Dániel Margócsy. 2022. "Datafication, Power and Control in Development: A Historical Perspective on the Perils and Longevity of Data." *Progress in Development Studies* :1-22.

Coletta, Claudio, and Rob Kitchin. 2017. "Algorhythmic Governance: Regulating the 'Heartbeat' of a City Using the Internet of Things." *Big Data & Society* 4(2): 1-16.

Cordella, Antonio, and Carla M. Bonina. 2012. "A Public Value Perspective for ICT Enabled Public Sector Reforms: A Theoretical Reflection." *Social Media in Government - Selections from the 12th Annual International Conference on Digital Government Research* 29(4):512–20.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. 2009. *Introduction to Algorithms*. 3rd ed. Cambridge and Malden, MA: MIT press.

Couldry, Nick. 2004. "Theorising Media as Practice." *Social Semiotics* 14(2):115–32.

Couldry, Nick. 2012. *Media, Society, World: Social Theory and Digital Media Practice*. Cambridge and Malden, MA: Polity Press.

Couldry, Nick. 2020. "Recovering critique in an age of datafication." *New Media & Society* 22(7):1135-1151.

Couldry, Nick, and Alison Powell. 2014. "Big Data from the Bottom Up." *Big Data & Society* 1(2):1-5

Coulthart, Stephen, and Ryan Riccucci. 2022. "Putting Big Data to Work in Government: The Case of the United States Border Patrol." *Public Administration Review* 82(2):280–89.

Crawford, Kate. 2021. The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale: Yale University Press.

Cresswell, John W, and Vicky L. Plano Clark. 2011. Designing and Conducting Mixed Methods Research. London: Sage Publications.

D'Agostino, Marcello, and Massimo Durante. 2018. "Introduction: The Governance of Algorithms." *Philosophy & Technology* 31(4):499–505.

Dalton, Craig M., and Jim Thatcher. 2014. "What Does A Critical Data Studies Look Like, And Why Do We Care?" https://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care

Danaher, John, Michael J. Hogan, Chris Noone, Rónán Kennedy, Anthony Behan, Aisling De Paor, Heike Felzmann, Muki Haklay, Su-Ming Khoo, John Morison, Maria Helen Murphy, Niall O'Brolchain, Burkhard Schafer, and Kalpana Shankar. 2017. "Algorithmic Governance: Developing a Research Agenda through the Power of Collective Intelligence." *Big Data & Society* 4(2):1-21.

DeLeon, Peter. 1999. "The Missing Link Revisited: Contemporary Implementation Research." *Review of Policy Research* 16(3-4):311–38.

Dencik, Lina. 2019. "Situating Practices in Datafication–from above and Below." Pp. 243–55 in *Citizen Media and Practice*, edited by H. C. Stephansen and E. Treré. Milton: Routledge. Routledge.

Dencik, Lina. 2020. "Mobilizing Media Studies in an Age of Datafication." *Television & New Media* 21(6):568–73.

Dencik, Lina. 2022. "The Datafied Welfare State: A Perspective from the UK." Pp. 145–65 in *New Perspectives in Critical Data Studies*, edited by A. Hepp, J. Jarke and L. Kramp. Palgrave Macmillan. https://link.springer.com/book/10.1007/978-3-030-96180-0#toc

Dencik, Lina, Arne Hintz, and Jonathan Cable. 2016. "Towards Data Justice? The Ambiguity of Anti-Surveillance Resistance in Political Activism." *Big Data & Society* 3(2):1-12.

Dencik, Lina, and Anne Kaun. 2020. "Datafication and the Welfare State." *Global Perspectives* 1(1):12912.

Dencik, Lina, Joanna Redden, Arne Hintz, and Harry Warne. 2019. "The 'golden View': Data-Driven Governance in the Scoring Society." *Internet Policy Review* 8(2):1–24.

Desrosières, Alain. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning*. Harvard: Harvard University Press.

Diakopoulos, Nicholas. 2014. "Algorithmic Accountability Reporting: On the Investigation of Black Boxes." *Two Center for Digital Journalism.* Report.

Diebold, Francis X. 2012. "On the Origin (s) and Development of the Term 'Big Data'." *PIER Working Paper No, 12-037*. https://ssrn.com/abstract=2152421

Dobrolyubova, Elena. 2021. "Measuring Outcomes of Digital Transformation in Public Administration: Literature Review and Possible Steps Forward." Journal of Public Administration and Policy 14(1):61–86.

Dunleavy, Patrick, Margetts, Helen, Bastow , Simon and Jane Tinkler. 2006. "New Public Management Is Dead—Long Live Digital-Era Government." Journal of public administration research and theory 16(3):467-494.

Dølvik, Jon Erik. 2013. "Grunnpilarene i de Nordiske Modellene." *Et Tilbakeblikk På Arbeidslivs-Og Velferdsregimenes Utvikling N*ordMod2030. Oslo: Fafo, Rapport 13.

Elish, Madeleine Clare, and Danah boyd. 2018. "Situating Methods in the Magic of Big Data and AI." Communication Monographs 85(1):57–80.

Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism.* Princeton: Princeton University Press.

Esteves, José and Rhoda C. Joseph. 2008. "A comprehensive framework for the assessment of eGovernment projects". Government information quarterly 25(1):118-132.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Feenberg, Andrew. 1991. *Critical Theory of Technology*. Vol. 5. New York: Oxford University Press.

Fimreite, Anne Lise, and Jan Erik Grindheim. 2007. *Offentlig Forvaltning.* Oslo: Universitetsforlaget.

Flensburg, Sofie, and Stine Lomborg. 2021. "Datafication Research: Mapping the Field for a Future Agenda." *New Media & Society.* Online First: 1-19.

Floridi, L., J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena. 2018. "AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28(4):689–707.

Frank, Lone. 2000. "When an Entire Country Is a Cohort." *Science* 287(5462):2398–99.

Fredriksson, Cecilia, Farooq Mubarak, Marja Tuohimaa, and Ming Zhan. 2017. "Big Data in the Public Sector: A Systematic Literature Review." *Scandinavian Journal of Public Administration* 21(3):39–62.

Fujimura, Joan. 2003. "Future Imaginaries" Pp. 176-199 in *Genetic nature/culture*, edited by A. H. Goodman, D. Heath, and M. S. Lindee. Berkeley: University of California Press.

Germundsson, Nora. 2022. "Promoting the Digital Future: The Construction of Digital Automation in Swedish Policy Discourse on Social Assistance." *Critical Policy Studies* Online First: 1–19.

Giest, Sarah N., and Bram Klievink. 2022. "More than a Digital System: How AI Is Changing the Role of Bureaucrats in Different Organizational Contexts." *Public Management Review* Online First: 1–20.

Gillespie, Tarleton. 2014a. "Algorithm [Draft][# Digitalkeywords]." *Culture Digitally*. https://culturedigitally.org/2014/06/algorithm-draft-digitalkeyword/

Gillespie, Tarleton. 2014b. "The Relevance of Algorithms." Pp.167-193 in *Media Technologies: Essays on Communication, Materiality, and Society* edited by T. Gillespie, P. Boczkowski, and K. Foot. Cambridge, MA: MIT Press

Gitelman, Lisa. 2013. *Raw Data Is an Oxymoron*. Cambridge, MA: MIT press.

Gorwa, Robert, Reuben Binns, and Christian Katzenbach. 2020. "Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance." *Big Data & Society* 7(1):1-15.

Gritsenko, Daria, Annette Markham, Holger Pötzsch, and Mariëlle Wijermars. 2022. “Algorithms, Contexts, Governance: An Introduction to the Special Issue.” *New Media & Society* 24(4):835–44..

Groves, Christopher, Karen Henwood, Fiona Shirani, Catherine Butler, Karen Parkhill and Nick Pidgeon. 2016. "The grit in the oyster: using energy biographies to question socio-technical imaginaries of ‘smartness’" *Journal of Responsible Innovation* 3(1):4-25.

Große-Bley, Jelena, and Genia Kostka. 2021. “Big Data Dreams and Reality in Shenzhen: An Investigation of Smart City Implementation in China.” *Big Data & Society* 8(2): 1-14.

Guenduez, Ali A., Tobias Mettler, and Kuno Schedler. 2020. “Technological Frames in Public Administration: What Do Public Managers Think of Big Data?” *Government Information Quarterly* 37(1):101406.

Habermas, Jürgen. 1970. "Technology and Science as 'Ideology'" Pp. 81-127 in Toward a Rational Society, edited by J. Habermas. Boston: Beacon Press

Hacking, Ian. 1986. “Making up People.” Pp.222-236 *Reconstructing Individualism,* edited by T. Heller and C. Brooke-Rose. Stanford: Stanford University Press.

Hagendorff, Thilo, and Katharina Wezel. 2020. “15 Challenges for AI: Or What AI (Currently) Can’t Do.” AI & Society 35(2):355–65.

Haraway, Donna. 1988. "Situated knowledges: the science question in feminism and the privilege of partial perspective." *Feminist Studies* 14(3):575-599.

Haverland, Markus, and Dvora Yanow. 2012. “A Hitchhiker’s Guide to the Public Administration Research Universe: Surviving Conversations on Methodologies and Methods.” *Public Administration Review* 72(3):401–8.

Heeks, Richard. 2006. *Implementing and Managing EGovernment: An International Text*. London: Sage.

Heeks, Richard, and Savita Bailur. 2007. “Analyzing E-Government Research: Perspectives, Philosophies, Theories, Methods, and Practice.” *Government Information Quarterly* 24(2):243–65.

Heidegger, Martin. 1977. *The Question Concerning Technology.* New York: Harper & Row

Hintz, Arne, Lina Dencik, and Karin Wahl-Jorgensen. 2018. *Digital Citizenship in a Datafied Society*. Cambridge, UK: John Wiley & Sons.

Hintz, Arne, Lina Dencik, Joanna Redden, Emiliano Treré, Jess Brand and Harry Warne. 2022. Civic Participation in the Datafied Society: Towards Democratic Auditing? Report. Data Justice Lab. https://datajusticelab.org/towards-democratic-auditing-civic-participation-in-the-scoring-society/

Hockenhull, Michael, and Marisa Leavitt Cohn. 2021. “Hot Air and Corporate Sociotechnical Imaginaries: Performing and Translating Digital Futures in the Danish Tech Scene.” *New Media & Society* 23(2):302–21.

Hoeyer, K. 2019. “Data as Promise: Reconfiguring Danish Public Health through Personalized Medicine.” *Social Studies of Science* 49(4):531–55.

Hong, Sounman, Sun Hyoung Kim, Youngrok Kim, and Jeongin Park. 2019. “Big Data and Government: Evidence of the Role of Big Data for Smart Cities.” *Big Data & Society* 6(1):1-11.

Hughes, Thomas Parke. 1993. *Networks of Power: Electrification in Western Society, 1880-1930*. Baltimore: JHU press.

Iliadis, Andrew, and Federica Russo. 2016. “Critical Data Studies: An Introduction.” *Big Data & Society* 3(2):1-7.

Introna, Lucas D. 2016. “Algorithms, Governance, and Governmentality: On Governing Academic Writing.” *Science, Technology, & Human Values* 41(1):17–49.

Issar, Shiv, and Aneesh Aneesh. 2022. “What Is Algorithmic Governance?” *Sociology Compass* 16(1): 1-14.

Jasanoff, Sheila. 2006. "Afterworld." Pp. 274-282 in *States of Knowledge: The Co-production of Science and Social Order*, edited by S. Jasanoff. New York: Routledge.

Jasanoff, Sheila, and Sang-Hyun Kim. 2009. “Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea.” *Minerva* 47(2):119–46.

Jasanoff, Sheila. 2015. “Imagined and Invented Worlds.” PP. 321-341 In Dreamscapes of Modernity. Sociotechnical Imaginaries and the Fabrication of Power, edited by S. Jasanoff and S.-H. Kim. Chicago: The University of Chicago Press.

Jasanoff, Sheila, and Sang-Hyun Kim. 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. Chicago, Ill: University of Chicago Press.

Jørgensen, Rikke Frank. 2021. “Data and Rights in the Digital Welfare State: The Case of Denmark.” *Information, Communication & Society* . Online first:1–16.

Just, Natascha, and Michael Latzer. 2017. “Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet.” *Media, Culture & Society* 39(2):238–58.

Katzenbach, Christian, and Lena Ulbricht. 2019. “Algorithmic Governance.” *Internet Policy Review* 8(4):1–18.

Keddell, Emily. 2015. “The Ethics of Predictive Risk Modelling in the Aotearoa/New Zealand Child Welfare Context: Child Abuse Prevention or Neo-Liberal Tool?” Critical Social Policy 35(1):69–88.

Kennedy, Helen. 2018. "Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication." Krisis: Journal for Contemporary Philosophy 1: 18-30.

Kennedy, Helen, Susan Oman, Mark Taylor, Jo Bates, and Robin Steedman. 2020. “Public Understanding and Perceptions of Data Practices: A Review of Existing Research.” Report. Living with Data, University of Sheffield.

Kim, Gang-Hoon, Silvana Trimi, and Ji-Hyong Chung. 2014. “Big-Data Applications in the Government Sector.” Commun. ACM 57(3):78–85.

Kirkpatrick, Graeme. 2017. Technology and Social Power. Houndmills: Palgrave Macmillan.

Kitchin, Rob. 2014a. “Big Data, New Epistemologies and Paradigm Shifts.” *Big Data & Society* 1(1):1-12.

Kitchin, Rob. 2014b. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Los Angeles: Sage Publications.

Kitchin, Rob. 2017. “Thinking Critically about and Researching Algorithms.” *Information, Communication & Society* 20(1):14–29.

Kitchin, Rob, and Tracey Lauriault. 2014. “Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work.” *The Programmable City Working Paper 2*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

Klievink, Bram, Bart-Jan Romijn, Scott Cunningham, and Hans de Bruijn. 2017. "Big Data in the Public Sector: Uncertainties and Readiness." *Information Systems Frontiers* 19(2):267–83.

Klinger, Ulrike, and Jakob Svensson. 2018. "The End of Media Logics? On Algorithms and Agency." *New Media & Society* 20(12):4653–70.

Kraemer, Kenneth, and John Leslie King. 2006. "Information Technology and Administrative Reform: Will E-Government Be Different?" *International Journal of Electronic Government Research (IJEGR)* 2(1):1–20.

Krick, Eva, and Cathrine Holst. 2021. "Governance by Hybrid Advisory Committees–A Hallmark of Social Democracy?" Pp. 113–30 in *Social Democracy in the 21st Century*, edited by D. E. Thorsen, Ø. Bratsberg and N. Brandal, Vol. 35. London: Emerald Publishing Limited.

Kuziemski, Maciej, and Gianluca Misuraca. 2020. "AI Governance in the Public Sector: Three Tales from the Frontiers of Automated Decision-Making in Democratic Settings." *Artificial Intelligence, Economy and Society* 44(6):1-13.

Larsson, S. 2018. "Algorithmic Governance and the Need for Consumer Empowerment in Data-Driven Markets." *Internet Policy Review* 7(2):1-13.

Lassiter, Luke Eric, and Luke Eric. 2005. *The Chicago Guide to Collaborative Ethnography*. Chicago: University of Chicago Press.

Lassiter, LukeEric. 2005. "Collaborative Ethnography and Public Anthropology." *Current Anthropology* 46(1):83–106.

Latour, Bruno and Steve Woolgar. 1979. Laboratory Life: The Social Construction of Facts. Beverly Hills: Sage Publications.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Milton Keynes: Open university press.

Law, John, 2004. *After method: Mess in social science research.* London: Routledge.

Lehtiniemi, Tuukka, and Minna Ruckenstein. 2018. "The Social Imaginaries of Data Activism." *Big Data & Society* 6(1):1-12.

Liu, Chuncheng, and Ross Graham. 2021. "Making Sense of Algorithms: Relational Perception of Contact Tracing and Risk Assessment during COVID-19." *Big Data & Society* 8(1):1-13.

Liu, Jun. 2022. "Social Data Governance: Towards a Definition and Model." *Big Data & Society* 9(2):1-14.

Liu, Shuhua Monica, and Qianli Yuan. 2015. "The Evolution of Information and Communication Technology in Public Administration." *Public Administration and Development* 35(2):140–51.

Löfgren, Karl, and C. William R. Webster. 2020. "The Value of Big Data in Government: The Case of 'Smart Cities.'" *Big Data & Society* 7(1):1-14.

Lomborg, Stine, Lina Dencik, and Hallvard Moe. 2020. "Methods for Datafication, Datafication of Methods: Introduction to the Special Issue." *European Journal of Communication* 35(3):203–12.

Lupton, Deborah. 2016. *The Quantified Self*. Cambridge, UK: John Wiley & Sons.

Lyon, David. 2014. "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique." *Big Data & Society* 1(2):1-13.

MacKenzie, Donald. 1996. *Knowing machines: Essays on technical change.* Cambridge: MIT Press

Maciejewski, Mariusz. 2017. "To Do More, Better, Faster and More Cheaply: Using Big Data in Public Administration." *International Review of Administrative Sciences* 83(1):120–35.

Mager, Astrid, and Christian Katzenbach. 2021. "Future Imaginaries in the Making and Governing of Digital Technology: Multiple, Contested, Commodified." New Media & Society 23(2):223–36.

Mantello, Peter. 2016. "The Machine That Ate Bad People: The Ontopolitics of the Precrime Assemblage." Big Data & Society 3(2):1-11.

Marcuse, Herbert. 1941. "Some social implications of modern technology." Zeitschrift für Sozialforschung 9(3): 414-439.

Marcuse, Herbert. 1964. One Dimensional Man: Studies in the Ideology of Advances Industrial Society. Boston: Beacon Press

Margetts, Helen and Patrick Dunleavy. 2013. "The second wave of digital-era governance: a quasi-paradigm for government on the Web." Philosophical Transactions of the Royal Society 371(1987):1-17.

May, Reuben A. Buford, and Mary Pattillo-McCoy. 2000. "Do You See What I See? Examining a Collaborative Ethnography." *Qualitative Inquiry* 6(1):65–87.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big data: A revolution that will transform how we live, work, and think.* London: John Murray.

McDonald III, Bruce D., Jeremy L. Hall, Janine O'Flynn, and Sandra van Thiel. 2022. "The Future of Public Administration Research: An Editor's Perspective." *Public Administration* 100(1):59–71.

McQuillan, Dan. 2016. "Algorithmic Paranoia and the Convivial Alternative." *Big Data & Society* 3(2):1-12.

McQuillan, Dan. 2018. "People's Councils for Ethical Machine Learning". Social Media and Society 4(2): 1-10.

Meijer, AJ, Manuel Pedro Rodríguez Bolívar, and J. Ramon Gil-Garcia. 2018. "From E-Government to Digital Era Governance and beyond: Lessons from 15 Years of Research into Information and Communications Technology in the Public Sector." *Journal of Public Administration Research and Theory:* 1–6.

Meijer, Albert. 2007. "Why Don't They Listen to Us? Reasserting the Role of ICT in Public Administration." *Information Polity* 12(4):233–42.

Meijer, Albert, and Victor Bekkers. 2015. "A Metatheory of E-Government: Creating Some Order in a Fragmented Research Field." Government Information Quarterly 32(3):237–45.

Meld. St. 27 2015-2016. 2016. Digital Agenda for Norge. Kommunal- og moderniseringsdepartementet. Government White Paper. https://www.regjeringen.no/no/dokumenter/meld.-st.-27-20152016/id2483795/

Meng, Amanda, and Carl DiSalvo. 2018. "Grassroots Resource Mobilization through Counter-Data Action." *Big Data & Society* 5(2):1-12.

Mergel, Ines, Noella Edelmann, and Nathalie Haug. 2019. "Defining Digital Transformation: Results from Expert Interviews." *Government Information Quarterly* 36(4):101385.

Mejias, Ulises A., and Nick Couldry. 2019. "Datafication." *Internet Policy Review* 8(4): 1-10.

Milan, Stefania. 2017. "Data Activism as the New Frontier of Media Activism." Pp. 151–63 in *Media activism in the digital age*, edited by V. Pickard and G. Yang. London: Routledge.

Milan, Stefania. 2020. "Techno-Solutionism and the Standard Human in the Making of the COVID-19 Pandemic." *Big Data & Society* 7(2):1-7.

Milan, Stefania, and Emiliano Treré. 2019. "Big Data from the South(s): Beyond Data Universalism." Television & New Media 20(4):319–35.

Milan, Stefania, and Lonneke van der Velden. 2016. " The Alternative Epistemologies of Data Activism". Digital Culture & Society 2(2):57-74.

Miles, Matthew B., and A. Michael Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. London: Sage Publishing.

Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society* 3(2):1-21.

Moats, David, and Nick Seaver. 2019. "'You Social Scientists Love Mind Games': Experimenting in the 'Divide' between Data Science and Critical Algorithm Studies." *Big Data & Society* 6(1):1-11.

Montgomery, John Dickey. 1967. *Sources of Administrative Reform: Problems of Power, Purpose, and Politics*. Bloomington: Harvard University Press.

Moon, M. Jae. 2002. "The Evolution of E-government among Municipalities: Rhetoric or Reality?" *Public Administration Review* 62(4):424–33.

Moses, Jonathan, and Torbjørn Knutsen. 2012. Ways of knowing: competing methodologies in social and political research. London: Red Globe Press.

Müller-Birn, Claudia, Leonhard Dobusch, and James D. Herbsleb. 2013. "Work-to-Rule: The Emergence of Algorithmic Governance in Wikipedia." Proceedings of the 6th International Conference on Communities and Technologies. Working paper.

Musiani, Francesca. 2013. "Governance by Algorithms." *Internet Policy Review* 2(3):1–8.

NESH. 2022. Forskningsetiske retningslinjer for samfunnsvitenskap og humaniora. https://www.forskningsetikk.no/retningslinjer/hum-sam/forskningsetiske-retningslinjer-for-samfunnsvitenskap-og-humaniora/

Neyland, Daniel. 2016. "Bearing Account-Able Witness to the Ethical Algorithmic System." *Science, Technology, & Human Values* 41(1):50–76.

Nikunen, Kaarina, and Jenni Hokka. 2020. "Welfare State Values and Public Service Media in the Era of Datafication." *Global Perspectives* 1(1):12906.

OED Online. 2022. "Actionable, Adj." *Oxford English Dictionary*.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Broadway books.

van Ooijen, Charlotte, Barbara Ubaldi, and Benjamin Welby. 2019. "A Data-Driven Public Sector: Enabling the Strategic Use of Data for Productive, Inclusive and Trustworthy Governance." *OECD Working Papers on Public Governance (33). Paris: OECD Publishing.*

O'Toole Jr, Laurence J. 2000. "Research on Policy Implementation: Assessment and Prospects." *Journal of Public Administration Research and Theory* 10(2):263–88.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard: Harvard University Press.

Passi, Samir, and Phoebe Sengers. 2020. “Making Data Science Systems Work.” *Big Data & Society* 7(2):1-13.

Pencheva, Irina, Marc Esteve, and Slava Jankin Mikhaylov. 2020. “Big Data and AI–A Transformational Shift for Government: So, What next for Research?” *Public Policy and Administration* 35(1):24–44.

Pinch, Trevor J., and Wiebe E. Bijker. 1984. “The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other.” *Social Studies of Science* 14(3):399–441.

Porter, Theodore M. 1996. “Trust in Numbers.” in *Trust in Numbers*. Princeton, New Jersey: Princeton University Press.

Ranerup, A., and HZ Henriksen. 2022. “Digital Discretion: Unpacking Human and Technological Agency in Automated Decision Making in Sweden’s Social Services.” *Social Science Computer Review* 40(2):445–61.

Redden, Joanna. 2018. “Democratic Governance in an Age of Datafication: Lessons from Mapping Government Discourses and Practices.” *Big Data & Society* 5(2):1-13.

Redden, Joanna, Lina Dencik, Arne Hintz, and Harry Warne. 2019. “Data Scores as Governance: Mapping and Analysing Changing Practices in the UK.” Report https://www.superscoring.de/2019/09/25/data-scores-as-governance-mapping-and-analysing-changing-practices-in-the-uk/

Redden, Joanna, Lina Dencik, and Harry Warne. 2020. “Datafied Child Welfare Services: Unpacking Politics, Economics and Power.” *Policy Studies* 41(5):507–26.

Reichertz, Jo. 2007. "Abduction: The logic of discovery of grounded theory. Forum: Qualitative Social Research." Pp. 214-228 in The SAGE handbook of grounded theory, edited by K. Charmaz and A. Bryant. London: Sage Publications.

Reutter, Lisa. 2018. “Unpacking the Socio-Technological Assemblage of Smart Algorithms - A Case Study on the Production of Machine Learning Algorithms in the Norwegian Labor and Welfare Administration.” Master's thesis. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2573466

Reutter, Lisa, and Hendrik Storstein Spilker. 2019. “The Quest for Workable Data. Building Machine Learning Algorithms from Public Sector Archives.” Pp. 95-107 in *The Democratization of Artificial Intelligence: Net Politics in the Era of Learning Algorithms*, edited by A. Sudmann. Berlin: transcript.

Rieder, Bernhard, and Guillaume Sire. 2014. “Conflicts of Interest and Incentives to Bias: A Microeconomic Critique of Google’s Tangled Position on the Web.” *New Media & Society* 16(2):195–211.

Rieder, Gernot. 2018. “Tracing Big Data Imaginaries through Public Policy: The Case of the European Commission.” Pp. 89–109 in *The Politics of Big Data*, edited by A. R. Sætnanm I. Schneider and N. Green. London: Routledge.

Rieder, Gernot, and Judith Simon. 2016. “Datatrust: Or, the Political Quest for Numerical Evidence and the Epistemologies of Big Data.” Big Data & Society 3(1):1-6.

Rinehart, Kerry Earl. 2020. "Abductive Analysis in Qualitative Inquiry." *Qualitative Inquiry* 27(2):303-311

Rogers, Everett M. 1995. “Lessons for Guidelines from the Diffusion of Innovations.” The Joint Commission Journal on Quality Improvement 21(7):324–28.

Rokkan, Stein. 1966. Norway: Numerical Democracy and Corporate Pluralism. Oslo: Chr. Michelsens Institutt.

Ruckenstein, Minna, and Natasha Dow Schüll. 2017. “The Datafication of Health.” Annual Review of Anthropology 46:261–78.

Ruppert, Evelyn. 2016. "Big data economies and ecologies." Pp. 13-29 in An End to the Crisis of Empirical Sociology? Trends and Challenges in Social Science Research, edited by L. Ryan and L. McKie. London: Routledge.

Ruppert, Evelyn. 2018. Sociotechnical Imaginaries of Different Data Futures. Dag van de Sociologie, Dutch and Flemish Sociological Association Conference and the Third Van Doorn Lecture. https://research.gold.ac.uk/id/eprint/23570/

Ruppert, Evelyn, Engin Isin, and Didier Bigo. 2017. "Data Politics." *Big Data & Society* 4(2):1-7.

Ryggvik, Helge. 2009. *Til siste dråpe. Om oljens politiske økonomi.* Oslo: Aschehoug.

Sadowski, Jathan. 2019. "When Data Is Capital: Datafication, Accumulation, and Extraction." *Big Data & Society* 6(1): 1-12.

Sadowski, Jathan, and Frank A. Pasquale. 2015. "The Spectrum of Control: A Social Theory of the Smart City." *First Monday* 20(7). https://firstmonday.org/article/view/5903/4660

Sánchez-Monedero, Javier, and Lina Dencik. 2019. "The Datafication of the Workplace." *Working Paper.* https://datajusticeproject.net/wp-content/uploads/sites/30/2019/05/Report-The-datafication-of-the-workplace.pdf

Sandvik, Kristin B. 2020. "'Smittestopp': If You Want Your Freedom Back, Download Now." *Big Data & Society* 7(2):1-11.

Sarker, Nazirul Islam, Min Wu, and Md Altab Hossin. 2018. "Smart governance through bigdata: Digital transformation of public agencies." 2018 international conference on artificial intelligence and big data (ICAIBD). IEEE: 62-70.

Saurwein, Florian, Natascha Just, and Michael Latzer. 2015. "Governance of Algorithms: Options and Limitations." *Info* 17(6):35-49.

Schartum, Dag Wiese. 2017. "Forvaltningsloven Og Elektronisk Forvaltning." *Lov Og Rett* 56(1):45–62.

Schatzki, Theodore R. 2011. "Where the Action Is (on Large Social Phenomena Such as Sociotechnical Regimes)." Sustainable Practices Research Group, Working Paper 1:1–31.

Schou, Jannik and Morten Hjelholt. 2018. "Digital state spaces: state rescaling and advanced digitalization." Territory, Politics, Governance 7(4):438-454.

Schiølin, Kasper 2020. "Revolutionary Dreams: Future Essentialism and the Sociotechnical Imaginary of the Fourth Industrial Revolution in Denmark." Social Studies of Science 50 (4): 542–566

Seaver, Nick. 2017. "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems." *Big Data & Society* 4(2):1-12.

Seaver, Nick. 2019. "Knowing Algorithms." Pp. 412–22. *DigitalSTS - A Field Guide for Science & Technology Studies,* edited by J. Vertesi and D. Ribes. Princeton University Press.

Shastri, Apoorva, and Mihir Deshpande. 2020. "A Review of Big Data and Its Applications in Healthcare and Public Sector." Pp. 55–66 in *Big Data Analytics in Healthcare*, edited by A. J. Kulkarni, P. Siarry, P. K. Singh, A. Abraham, M. Zhang, A. Zomaya, and F. Baki. Cham: Springer International Publishing.

Shouran, Zaied, TK Priyambodo, and Nur Rokhman. 2019. "EGovernment Transformation: Literature Review." *International Journal of Scientific & Technology Research* 8(06):208–12.

Silverstone, Roger, and Leslie Haddon. 1996. *Design and the Domestication of Information and Communication Technologies: Technical Change and Everyday Life*.Oxford: Oxford University Press.

Smith, GJD. 2020. "The Politics of Algorithmic Governance in the Black Box City." *Big Data & Society* 7(2):1-9.

Star, Susan Leigh. 1988. "Introduction: The Sociology of Science and Technology." *Social Problems* 35(3):197–205.

Star, Susan Leigh, and James R. Griesemer. 1989. "Institutional Ecology,Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19(3):387–420.

Stephansen, Hilde C., and Emiliano Treré. 2019. Citizen Media and Practice: Currents, Connections, Challenges. Milton: Routledge.

Storm Pedersen, John, and Adrian Wilkinson. 2018. "The digital society and provision of welfare services." International Journal of Sociology and Social Policy 38(3/4): 194-209.

Swidler, Ann. 2001. "What Anchors Cultural Practices" Pp. 74-92 in The Practice Turn in Contemporary Theory, edited by T. Schatzki, K. Knorr Cetina and E. von Savigny. London: Routledge

Tavory, Iddo, and Stefan Timmermans. 2014. Abductive Analysis: Theorizing Qualitative Research. Chicago, Ill: University of Chicago Press.

Taylor, Charles. 2004. *Modern social imaginaries.* Durham: Duke University Press.

Taylor, Linnet. 2017. "What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally." *Big Data & Society* 4(2):1-14.

Thue, Lars. 2003. For egen kraft. Kraftkommunene og det norske kraftregimet 1887–2003. Oslo: Abstrakt.

Timmermans, Stefan, and Iddo Tavory. 2022. *Data Analysis in Qualitative Research: Theorizing with Abductive Analysis*. Chicago, Ill: University of Chicago Press.

Tracy, Sarah J. 2010. "Qualitative Quality: Eight "Big-Tent" Criteria for Excellent Qualitative Research." Qualitative Inquiry 16(10):837-851.

Tupasela, Aaro, Karoliina Snell, and Heta Tarkkala. 2020. "The Nordic Data Imaginary." *Big Data & Society* 7(1):1-13.

Turow, Joseph, and Nick Couldry. 2018. "Media as Data Extraction: Towards a New Map of a Transformed Communications Field." *Journal of Communication* 68(2):415–23.

Van Dijck, José. 2014. "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology." *Surveillance & Society* 12(2):197–208.

Veale, Michael, Max Van Kleek, and Reuben Binns. 2018. "Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making." Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.

Velkova, Julia, and Anne Kaun. 2021. "Algorithmic Resistance: Media Practices and the Politics of Repair." *Information, Communication & Society* 24(4):523–40.

Wagner, B. 2019. "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems." *Policy and Internet* 11(1):104–122.

Wajcman, Judy. 2015. Pressed for Time The Acceleration of Life in Digital Capitalism. Chicago: The University of Chicago Press.

Wang, Pei. 2019. "On Defining Artificial Intelligence." *Journal of Artificial General Intelligence* 10(2):1–37.

Wijermars, Mariëlle, and Mykola Makhortykh. 2022. "Sociotechnical Imaginaries of Algorithmic Governance in EU Policy on Online Disinformation and FinTech." *New Media & Society* 24(4):942–63.

Winner, Langdon. 1993. "Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology." *Science, Technology, & Human Values* 18(3):362–78.

Wirtz, Bernd W., Paul F. Langer, and Carolina Fenner. 2021. "Artificial Intelligence in the Public Sector - a Research Agenda." *International Journal of Public Administration* 44(13):1103–28.

Woolgar, Steve. 1990. "Configuring the User: The Case of Usability Trials." *The Sociological Review* 38(1l):58–99.

Yang, Kaifeng, and Gerald J. Miller. 2007. Handbook of Research Methods in Public Administration. Boca Taron: Taylor & Francis

Yeung, Karen. 2017. "'Hypernudge': Big Data as a Mode of Regulation by Design." Information, Communication & Society 20(1):118–36.

Yeung, Karen. 2020. "From new public management to new public analytics." ADM workshop 1: conceptualizing automation in the welfare sector, November 16 2020.

Zarsky, Tal. 2016. "The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making." *Science, Technology, & Human Values* 41(1):118–32.

Zejnilović, Leid, Susana Lavado, Íñigo Martínez de Rituerto de Troya, Samantha Sim, and Andrew Bell. 2020. "Algorithmic Long-Term Unemployment Risk Assessment in Use: Counselors' Perceptions and Use Practices." *Global Perspectives* 1(1):12908.

Ziewitz, Malte. 2016. "Governing Algorithms: Myth, Mess, and Methods." *Science, Technology, & Human Values* 41(1):3–16.

van Zoonen, Liesbet. 2020. "Data Governance and Citizen Participation in the Digital Welfare State." *Data & Policy* 2(e10).

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019*. New York: PublicAffairs.

Zwitter, Andrej. 2014. "Big Data Ethics." *Big Data & Society* 1(2):1-6.

[Paper 1]
Constraining Context: Situating Datafication in Public Administration

This paper is published in *New Media & Society* 24(4):903-921.

Special Issue Article

new media & society
2022, Vol. 24(4) 903–921

DOI: 10.1177/14614448221079029
journals.sagepub.com/home/nms

Constraining context: Situating datafication in public administration

Lisa Reutter
Norwegian University of Science and Technology, Norway

Abstract

The imaginary of data-driven public administration promises a more effective and knowing public sector. At the same time, corporate practices of datafication are often hidden behind closed doors. Critical algorithm studies, therefore, struggle to access and explore these practices, to produce situated accounts of datafication and possible entry points to reconfigure the emerging data-driven society. This article offers a unique empirical account of the inner workings of data-driven public administration, asking the overall question of how sociotechnical imaginaries of datafication are constrained in the context of public administration. Teams working on datafication in two Norwegian public sector entities have been followed and interviewed over the course of 2 years (2018–2020). While sociotechnical imaginaries thrive in organizational culture and policy discourse alike, the observed data teams struggle to realize data assemblages due to a variety of structural and institutional constraints.

Keywords

Constraints, critical algorithm studies, data assemblage, datafication, public administration, sociotechnical imaginaries

Introduction

Across European countries, interest in the sociotechnical imaginary of data-driven public administration is growing (Misuraca and van Noordt, 2020). According to the Organisation for Economic Co-operation and Development (OECD), data-driven technology promises

Corresponding author:
Lisa Reutter, Department of Sociology and Political Science, Norwegian University of Science and Technology, Postboks 8900, Torgarden, Trondheim 7491, Norway.
Email: lisa.m.reutter@ntnu.no

to deliver more effective state benefits to families in need, save the lives of patients, protect children from abuse, and predict possible terrorist attacks (van Ooijen et al., 2019)—in other words, improving both service delivery and the efficiency of the public sector. While sociotechnical imaginaries represent collectively held and institutionally stabilized future visions of society, they nevertheless can be multiplied and contested (Mager and Katzenbach, 2021). Furthermore, future visions of datafication are mediated and negotiated by the institutions, workers, and users set on the quest to materialize the grand ideas of datafication (Caplan et al., 2020). This article asks the overall question: how does the context of public administration constrain the sociotechnical imaginaries of datafication as these are realized in specific data assemblages? The article offers unique empirical insight into the inner workings of public administration datafication in Norway and contributes to critical data and algorithm studies by situating datafication beyond Internet platforms, providing possible entry points for the reconfiguration of the data-driven society.

Data-driven technology is far from benign. Social scientists have pointed out a variety of shortcomings tied to the ever-growing influence of datafication in society, such as its impenetrable opaqueness, reinforcement of discrimination, and facilitation of surveillance (boyd and Crawford, 2012; Pasquale, 2015; van Dijck, 2014; Zuboff, 2019). A growing body of case studies has begun to address the unintended consequences of data-driven systems in public administration (Allhutter et al., 2020; Eubanks, 2018). Ultimately, this critique addresses a shift in the power dynamics between governments and citizens, as the state's potential to understand, predict, and control the activities of citizens is vastly enhanced (Dencik, Hintz, et al., 2019). Data activist studies have, therefore, started to research bottom-up initiatives in response to unjust data practices (Beraldo and Milan, 2019). This article builds on this research and directs analytical attention toward the ways dominant imaginaries of datafication are mediated and responded to within various contexts in which they are employed, prior to their working in the world (Kazansky and Milan, 2021). By doing so, the article aims to produce an actionable critique rather than a general and oversimplified account of datafication. This critique takes into consideration the variety of conditions of production and implementation and how these might mediate the power and influence of data-driven technology. Recognizing the already existing mediations helps to extend the understanding of how to respond to and govern datafication, to mitigate data harms, and to facilitate discussion about the ways in which datafication is contingent and therefore might be otherwise.

Social science researchers must find a balance between problematizing general trends in society and particular applications (Dencik, Redden, et al., 2019). Internet platform datafication has received significant scholarly interest (Bucher, 2018; Mejias and Couldry, 2019). Corporate practices of datafication, however, are often hidden behind closed doors. Researchers, therefore, struggle to access and explore datafication practices and to establish a balance between myth and hype and actual applications of data-driven systems (Elish and boyd, 2018; Ziewitz, 2015). Research carries the risk of disregarding how imaginaries are negotiated by the institutions, workers, and users as they strive to realize these future visions, and researchers may overdetermine its power or fetishize data and algorithms in this regard (Caplan et al., 2020; Thomas et al., 2018). Earlier research on administrative reforms related to data-driven public administration has already shown that the public sector encounters a variety of

obstacles when negotiating such visions of the future (Andrews, 2019; Mergel et al., 2016; Redden, 2018). Observed obstacles are often tied to the very nature of public administration and its constituent institution. Public administration consists of thousands of human minds; the work of its organizational units is heavily governed by legal documents, organizational structures, and existing data infrastructures. Its subjects are not consumers, but citizens; its goals are decided by politicians rather than corporate managers. Its inner workings are not "black boxed" because of trade secrets, but instead open for scrutiny by law.

The case studies presented in this article were conducted in two Norwegian public sector entities and use machine learning projects as an entry point into newly established data assemblages. The Norwegian Tax Administration and the Norwegian Labor and Welfare Administration are the two largest public sector entities in Norway, owning and managing vast amounts of data on its citizens and entrusted with broad public mandates. Data teams have been followed and interviewed over the course of 2 years (2018–2020). Before presenting the analysis of public administration mediations, the article introduces the concepts of datafication, sociotechnical imaginaries, and data assemblages, along with its research design.

Approaching datafication: from sociotechnical imaginaries to data assemblages

Datafication is the practice of quantifying every aspect of life so that it can be analyzed (Mayer- Schönberger and Cukier, 2013). The datafication of public administration consists of two interwoven processes: the use of *more and different data* (big government data) and the deployment of *more advanced methods* to analyze these data and feed it back into existing work processes. These interwoven practices materialize in concrete projects of automated decision-making, sorting algorithms, and the development of decision support tools enabled by machine learning. Datafication represents a strategy of administrative reform primarily concerned with the reform of knowledge production practices and the introduction of algorithmic forms of ordering (Yeung, 2020). Data-driven public administration operates according to the logics of categorization, classification, scoring, and selection and presents an extension and profound change in public sector data practices (Dencik et al., 2019). The datafication of public administration is built on the ideological foundation of *dataism*, the belief that data can promote a better, more effective, and more 'objective' societal apparatus (van Dijk, 2014). It is this belief in data's potential that is the core of the sociotechnical imaginary of data-driven public administration.

The concept of sociotechnical imaginaries addresses the normative function of "collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology" (Jasanoff and Kim, 2015: 4). This analytical tool addresses the co-production of technology, society, and politics and provides researchers with a framework within which to investigate how practitioners and institutions make sense of technological change. Sociotechnical imaginaries provide visions of the kinds of society that sociotechnical

change could bring into being; they shape what is thinkable and the practices through which actors perform their roles (Ruppert, 2018). Connected to social and technological order, they are, therefore, often associated with state power and the selection of priorities, allocation of funds, and investment in material infrastructures (Jasanoff and Kim, 2009). Datafication lies at the heart of the future imaginary of the welfare state and guides and informs public administration reforms (Dencik and Kaun, 2020). The data-driven imaginary is built on the idea of fast and cheap data analysis, the possibility of gleaning data from entire populations, data speaking for itself, and agnostic algorithms that guarantee an impartial view from nowhere (Rieder and Simon, 2016). Such sociotechnical imaginaries include notions of proactive, rather than reactive, modes of governance. Datafication provides the public sector with a sense of being able to do more, better, faster, and more cheaply and is therefore perceived as a solution to the growing complexity of society and administration and as a tool to reduce uncertainty (Maciejewski, 2017; Strauß, 2015). The vast amounts of data in Nordic registries were into this overall sociotechnical imaginary of the public sector as a data gold mine (Tupasela et al., 2020).

The content and performativity of sociotechnical imaginaries of datafication has been discussed and assessed in a variety of ways in recent years (see, for example, Lehtiniemi and Ruckenstein, 2019; Tupasela et al., 2020; Williamson, 2018). Although studying sociotechnical imaginaries themselves presents interesting insights into the wider origin and legitimation of data-driven technology in public administration, it provides little insight into the ways they reconfigure public administration and are themselves altered in concrete settings. Jasanoff and Kim (2015) therefore emphasize the need to investigate how sociotechnical imaginaries enter, are translated, and negotiated into assemblages of materiality, which this article intends to do.

Kitchin and Lauriault's (2014) framework of the data assemblage holds the potential to advance this understanding further. This framework is intended to illuminate the contextual and situational forces that influence the way data systems are shaped, to gain a better appreciation of the work they do in the world (Kitchin, 2014: 24). Data assemblages consist of a variety of apparatuses (entwined and contingent parts of the larger system) and elements (see Table 1).

These elements work together to produce a data assemblage not only discursively but also materially. Deconstructing the assemblage of data-driven technology allows researchers to show that these technologies are value-laden, socially constructed entities influenced by a variety of technological, political, social, and economic apparatuses (Kitchin, 2017). In their analysis of datafied child welfare services, Redden et al. (2020) apply this framework and show how systems of thought, ownership structures, policy agendas, organizational practices, and legal frameworks influence the datafication of public administration. They use these influencing factors to point out the sociotechnical aspects of datafication and the value-ladenness of systems. The following analysis is an attempt to reframe this argument: rather than regarding the apparatuses presented by Kitchin and Lauriault (2014) only as value-laden influencing factors, this analysis attempts to regard them as possible mediators that constrain sociotechnical imaginaries and allow for reconfigurations as these become specific data assemblages. This builds on the idea advocated by Tanweer et al. (2016) and Redden and Grant (2020) to study points

Table 1. The apparatus and elements of a data assemblage (Kitchin, 2014).

Apparatus	Elements
Systems of thought	Modes of thinking, philosophies, theories, models, ideologies, rationalities, etc..
Forms of knowledge	Research texts, manuals, magazines, websites, experiences, word of mouth, chat forums, etc..
Finance	Business models, investment, venture capital, grants, philanthropy, profit, etc..
Political economy	Policy, tax regimes, public and political opinion, ethical considerations, etc..
Governmentalities and legalities	Data standards, file formats, system requirements, protocols, regulations, laws, licensing, intellectual property regimes, etc..
Materialities and infrastructures	Paper/pens, computers, digital devices, sensors, scanners, databases, networks, servers, etc..
Practices	Techniques, ways of doing, learned behavior, scientific conventions, etc..
Organizations and institutions	Archives, corporations, consultants, retailers, government agencies, universities, conferences, clubs and societies, committees and boards, communities of practice, etc..
Subjectivities and communities	Of data producers, curators, managers, analysts, scientists, politicians, users, citizens, etc..
Places	Laboratories, offices, field sites, data centers, etc., and their agglomerations
Marketplace	For data, its derivates, analysist, analytical software, interpretations, etc..

of breakdown, where progress is stopped due to material and discursive limitations; these provide valuable insights and allow new imaginations and configurations of data assemblages to be developed. Bridging the concepts of sociotechnical imaginaries and data assemblages, this article shows how collectively held future visions are translated and negotiated into assemblages of materiality. This also creates the potential to find an analytical balance between problematizing general trends in society and emphasizing particular applications (Dencik et al., 2019). The sociotechnical imaginary of the data-driven society can be observed across both private and public institutions, but the apparatuses of data assemblages are highly context-dependent.

Researching public administration datafication

Machine learning, where algorithms are fed with big government data, is only one of the technologies associated with the datafication of public administration. This article uses this to illustrate the realization of sociotechnical imaginaries of data-driven public administration into specific data assemblages. Machine learning projects present an entry point into the data assemblage of data-driven technology; this article offers a concrete example of how such administrative reforms are (or are not) implemented.

Kitchin (2017) and Dencik (2019) encourage research into specific datafication projects and their contexts. This article therefore employs an empirical research design consisting of fieldwork, interviews, meeting observations, and document study. By law, the Norwegian public sector is required to provide access for researchers, which provides unique insight into its inner workings.[1] Two data teams in two different public entities have been followed over 2 years, the Norwegian Labour and Welfare Administration (NAV) and the Norwegian Tax administration (SKATT). Neither entity relies on private sector competence, which further eases access to the field itself. Contact with the NAV team was established in 2017 and with SKATT at an AI fagforum meeting in 2018. As the progression of these projects is slow, the NAV team was followed from January 2018 to May 2019 and the SKATT team from June 2019 to November 2019. This included five visits to each of these public entities, with each visit lasting 1 week or less. In total, the assembled data include 37 days of field notes, 17 interviews with key organizational actors and interviews with the data team, 16 meeting observations, 12 PowerPoint presentations, and approximately 60 pages of strategic documents or internal documentation. Several check-in meetings were conducted after the initial field work to follow-up on the projects. To understand the wider foundations of the work more thoroughly, I also attended and observed several industry and public sector conferences in 2018 and 2019. At conferences, notes were taken both at presentations and reflecting informal discussions among data team members. Key policy documents were analyzed: these included the digitalization strategy for the public sector, the Norwegian AI strategy, the concept selection study on sharing of data within the public sector, and mandate letters for the two organizations (2017–2019). An abductive analysis of the data material was conducted (as described by Tavory and Timmermans, 2014), based on in-depth reading of literature, immersion in fieldwork, and subsequent coding in NVivo. The element of surprise—the fact that most observed datafication efforts seem to meet a variety of obstacles along the way despite strong sociotechnical imaginaries—guided this research project significantly.

Research site: the NAV data team

NAV is the largest Norwegian public agency and is at the forefront of an ongoing nationwide digital transformation. It is a public welfare agency that delivers over 60 benefits and services, such as unemployment benefits and pensions. Operating under the Ministry of Labour, NAV has approximately 19,000 employees and manages approximately one-third of the overall state budget. NAV established a "big data"-laboratory in 2016, collecting a variety of actors from across the organization and exploring possibilities for big government data recirculation. The NAV data team resulted from this initiative, as part of the newly established division for Data and Insight. This section is mandated to collect all environments that develop and manage data products in the organization and is placed within the IT department. The NAV data team has now approximately 10 data workers from a variety of backgrounds, including physics, political science, and data engineering. At the time of writing, NAV has no machine learning model in production (November 2020).

One of the major projects followed in the fieldwork was the "sickness benefit" project. This project was established in 2018 and aims to have a first pilot in production by

December 2020. The project intends to predict the sickness absence of employees and give an indication for the necessity of an assessment meeting between a NAV case worker, the employee on sick leave, and the employer. The project has two full-time data team members and trains its models with sick leave certificates, using a boosting tree model. At this time, the final product is intended to be a decision support tool in the expert system for the case workers.

Research site: the SKATT data team

SKATT operates under the Ministry of Finance. Its main objective is to secure the funding of the welfare state through tax collection. It employs approximately 6500 people and has a variety of control and tax functions. The SKATT data group has a well-established team of data analysists, having existed since the early 2000s. Its main tasks include both traditional data analysis and report production, and project-based recirculation of data. This team has 27 employees based in two locations, and most of these employees still work on traditional analysis. SKATT has seven models in production that can be classified as machine learning.

One of the main projects followed in the field is the value-added tax project (PSU project). This project was established around 2013, aiming to systemize and automate the selection of control cases in value-added tax reporting by companies. Initially a rule-based engine (using if-then selection criteria) was employed in this process, but the organization's control unit wished to further optimize this process. Machine learning was not within the initial mandate of the data team's work but was explored along the way. The developed PSU model, a supervised machine learning model, is now one out of the seven working machine learning models in production within the agency. Based on historical data, it risk-scores and sorts all incoming reports, which are then controlled by street-level bureaucrats. Controlled and resolved cases are then fed back into the machine learning model.

Constraining context: toward a data-driven Norwegian public administration

Although the Norwegian public sector is one of the most digitized in Europe and sociotechnical imaginaries are enacted widely, a recent general mapping shows that there are few data-driven technologies in production (Broomfield and Reutter, 2019). Both data teams observed in this study have struggled to develop and implement new data assemblages within their organizations. It therefore appears that the context of Norwegian public administration is one where sociotechnical imaginaries flourish within the organizational discourse but where new data assemblages struggle to stabilize, despite the vast amounts of registry and case-work data on citizens. The following analysis reflects an attempt to deconstruct data-driven technology production and implementation within Norwegian public administration. It follows the sociotechnical imaginary from collectively held visions of the future as they do or do not stabilize into concrete data assemblages. This analysis begins by presenting the policy and the strategy guiding the work of data teams; then it moves on to the organizational and institutional constraints that

provide the framework within which the data teams work, the legal and regulatory environment for data utilization, the data infrastructure, and the practical aspects of implementing machine learning. The apparatuses presented in the analysis do not offer an exhaustive list of obstacles that mediate data-driven technologies in the process of their production. They serve, however, as a starting point to engage with these administrative reforms, to better understand how future visions are negotiated and translated into public administration practices.

Policy and strategy

Sociotechnical imaginaries are often associated with the ways nation states and governmental actors envision technological development and unfold their power to imagine, govern, and program technologies (Jasanoff and Kim, 2009). This is strongly evident in the field of public administration. Systems of thought and sociotechnical imaginaries about data-driven public administration are interwoven, manifested, circulated, and enacted in both national and multinational policy and strategy documents, as, for example, shown in the introduction of the national AI strategy:

> We know that Norway will be affected by an aging population, climate change and increased globalization, and that we must work smarter and more efficiently to maintain competitiveness and the level of welfare in the years to come. Digitalisation and new technology are the key to achieve this—and artificial intelligence will be central. (Ministry of Local Government and Modernisation, 2020)

The sociotechnical imaginary of public sector datafication in Norway promises to solve the problem of ineffective public administration. As do other countries, Norway supposedly faces a variety of future challenges, any of which might decrease the level of national welfare. These challenges include the downsizing of the oil sector, growing immigration, climate change, and globalization. Norwegian value creation in both public and private entities is deemed to be dependent on a thriving data-driven economy and government, according to policy documents. Datafication is thus regarded a common sense solution and part of the evolution of public administration. Public administration is concerned with the implementation of government policy and the administration and execution of public law. Policy documents, therefore, reduce sociotechnical imaginaries to actionable plans and focus areas.

Policy documents stress the tremendous unused potential in Norway's data sources. Much emphasis is placed on the sharing and re-use of data in public administration and between public and private organizations. The sheer availability of data is expected to lead to new solutions and improved services. This vision of the future is increasingly influenced by the private sector propagating notions of technological progress and benefit, as evident in a public spheres elsewhere (Hockenhull and Cohn, 2021). This could be observed in relation to the stakeholders consulted in policy production and the conferences visited in the field work. The private sector regards data as a key value-creation opportunity. According to national policy, users of public services also experience the services provided by the public sector as bothersome. Data sharing is expected to enable new data practices and more effective and seamless public service.

Policy documents are an important apparatus within the data assemblage, as they frame what and how data are used and recirculated (Kitchin, 2014). They provide public sector leadership with legitimation and a planning framework for the implementation of administrative reforms. Mandate letters, given to the public entities observed, shape the actual day-to-day work in the public entities. These are produced every year by the governing ministry and determine what the public entity is to work on through a 1-year period. They provide fiscal frameworks, priorities, expected results, and reporting requirements. The areas of special interest within these documents often define where the data teams will work and what kind of projects are initiated. Sickness benefit follow-up, for example, has been key to several mandate letters.

Organizational scope

Organizations and institutions are key apparatuses of the data assemblage (Kitchin, 2014). In addition to mandate letters, each of the organizations has detailed strategies which lay out long-term plans for the organization. These offer descriptions of how the organization is expected to operate in the future, echoing the sociotechnical imaginaries outlined at policy level and emphasizing the tremendous potential for effective change and the fear of missing out on a technological opportunity.

> Technological developments can make the agency more efficient and free up resources for other tasks. [. . .] It is likely that several of the agency's tasks over time, in whole or in part, will be replaced by smart ICT solutions. The Tax Administration must avoid ending up in a situation where the agency's systems and solutions become outdated and incompatible with the rapidly changing outside world. (Skatteetaten 2025, 2017)

The institutional framework of public administration, therefore, is an important factor shaping the data assemblage. Data teams act as executive agents of sociotechnical imaginaries within their organizations and play a key role in maintaining and circulating these imaginaries. Within a public sector organization, however, the data team's influence is mediated by its assigned scope. Both of the data teams observed are relatively small compared to the size of their organizations. The teams are part of a larger effort to recirculate data with stakeholders across the whole organization. Meanwhile, the two teams are placed in different sections in their organizations. The SKATT data team is relatively well-established within the analysis department, having its origins in the data warehouse section established in the early 2000s. The NAV data team, however, was created recently and is part of the IT department. The placement of a data team and its size, in addition to its organizational constraints, limits what it can do. Tight connections to the IT department ease the integration of systems into already existing digital infrastructure, as observed in NAV. At the same time, extensive experience curating, accessing, and analyzing data provide the SKATT data team with competence in the field of data analysis and existing data infrastructures.

Most datafication efforts and projects are led by the central core of public administration and are intended to be spread to lower levels. The public entities examined in this study differ in size, organization, and administrative scope. The NAV data team has a

wider administrative scope and a more fragmented and distributed organizational structure than the SKATT data team. This has complicated data access, due to differences in legal mandates between the municipal entities and the central level at NAV, according to member of the data team.

The financial apparatus of Norwegian public administration places great emphasis on project-based practices. Projects are mostly owned and ordered by other entities within the organization. Data teams are expected to deliver projects to these entities, which manage and decide how much to spend. As projects are time-limited, and data-driven efforts are often costly and regarded as long-term strategies by these organizations, several projects were ''*put on ice*'' during the period of observation because they either lacked resources or were not prioritized by the project owner. The data teams do not initiate projects by themselves, but instead are expected to act on "needs" identified by others within the organization. Surprisingly, both teams experienced little interest in data-driven technologies in their wider organizational environment. The SKATT team has struggled to identify possible future projects and, on multiple occasions, had to initiate an internal effort to find projects. The NAV data team spent its first months in existence presenting what machine learning is, and how it can be used, to a variety of actors within the organization. Both data teams highlighted that the so-called "need" for data-driven technology often needs to be produced by the data team together with the domain knowledge experts. Presentations about what data-driven technology is, and what it can help to accomplish, seemed to be a key activity for the circulation and enactment of sociotechnical imaginaries. The data teams use significant resources to distribute and circulate these visions of desirable futures for data-driven public administration, building authenticity through pilot projects or success stories. This has been especially difficult for the NAV data team, as they do not yet have a success story of their own. As Hockenhull and Cohn (2021: 16) point out in their study on corporate sociotechnical imaginaries of datafication, these imaginaries work in complex ways, being both vacuous and productive. These authors use the notion of "hot air" to point out the way sociotechnical imaginaries can reinforce common sense notions of progress, using buzz words to create a rhetorical ethos and building authority through exemplars.

Overall frameworks for projects are decided on by the domain experts as project owners. The framework of the sickness benefit project and its goals were assigned to the data team members, with the specific goal being to predict the necessity of assessment meetings. What the data team members actually attempted to predict is whether a worker would be on sick leave for longer than 16 weeks, which according to subject matter experts indicates a need for follow-up. What can be done, and how, is therefore often quite restricted and rule-based and is not steered by data as such. As one of the data team members points out: "*The 16 weeks are maybe not an ideal cut. But we make the best out of it*" (NAV data team member).

The project-based organization of work also leads to a high degree of dependence on other actors and sections within the organization. The data teams' progress, therefore, also depends on their ability to cooperate across the organization and to include key actors in their work: this includes subject matter experts, data managers, and legal staff. A key apparatus of the data assemblage is the form of knowledge itself. The public mandate of NAV is often connected to relatively vulnerable situations in people's lives, such

as unemployment, pensions, and illness; the organization therefore regards social work as a key form of knowledge. SKATT, on the contrary, regards accounting as its key form of knowledge, which is often connected to the practice of control.

Legal mandates

Legal frameworks are regarded as inhibiting the progress of datafication by data team members and have therefore received significant attention in recent policy documents. The translation and negotiation of sociotechnical imaginaries into material assemblages depends on an apparatus of governmentalities and legalities (Kitchin, 2014). Public administration represents the executive sphere of politics in modern democracies, and its realm of action is tied to some legal frameworks and mandates. In the Norwegian context, the public administration act, the context specific NAV law, and the tax administration act are each central to how specific public services are regulated, encompassing everything from what services are provided to how they should be performed. The analysis arising from this study clearly shows that mediation through law heavily influences new data assemblages in both data teams; this apparatus also manages to stop or delay most projects in these organizations.

"*Legal assessments are a real show-stopper*" (Project lead, SKATT). These data teams experience two main areas as especially challenging: (1) legal mandates that prohibit the recirculation of data collected in one area of the public entity into other areas of practice, and (2) the lack of legal frameworks to regulate the linking of different data sources. This concerns both existing legal frameworks and the absence of frameworks. Where legal frameworks do not yet exist juridical decisions can be restrictive, stopping projects before they start or along the way. Legal staff in both organizations struggle with the concept of government data recirculation. For example, the SKATT data team repeatedly noted that the assessment of projects can be frustrating and unpredictable and, in some cases, is experienced as essentially random. Existing legal frameworks significantly constrain projects in both data teams, for example, by excluding specific variables or datasets from use or analysis.

All data projects must be tied to a legal mandate, meaning that data teams and subject matter experts spend a significant amount of time identifying sections in the public administration act that can be connected to their projects, to legitimate the use of data-driven technology. In the case of the NAV sickness benefit project, the folketrygdloven, Chapter 8, § 8–7 a states that *NAV must hold a meeting together with the sick employee, the employer, before the illness benefit reaches 26 weeks. This is unless a meeting is considered to be clearly unnecessary.* Here, the sentence "considered to be clearly unnecessary" has been used to legitimize the whole project. The machine learning model is supposed to show what is "clearly unnecessary" by predicting cases that do not require follow-up. This decision was previously taken by the case worker responsible for sickness benefit follow-up, independent of any data analysis. It is now intended to be taken by case workers informed by the decision support tool the team has developed. As both the SKATT and NAV data teams point out, the legal mandates tied to control issues are often wider. It is therefore easier to get projects tied to control approved.

Most machine learning work at SKATT begins with simple logistic regressions, as nearly all projects are built on already existing rule-based automated sorting. The Norwegian Public Administration Act clearly states that decisions must be explainable and transparent (Public Administration Act, 2021). The data teams therefore perceive that the range of methods allowed is small. Most of the models developed and discussed at the point of observation were decision tree models or advanced logistic regressions. Supervised machine learning is often chosen rather than complex deep learning or neural networks: the latter are not explainable and therefore cannot be used, even if they score higher for accuracy in the test data. The NAV data team, for example, has developed a neural network to analyze the text in the certificates but decided to stop as this would require further legal assessments. It is often not the coding work itself that is negotiated, but its outcomes and foundations. Both teams spent a significant amount of time assessing the question of fairness in their machine learning models, asking questions, such as what is fairness in machine learning models in general? What kind of fairness measures should be used in which situation, and who can decide on this? The NAV team in particular has worked both methodologically and practically on this issue.

While the SKATT data team has obtained approval for several projects, the NAV data team has often been blocked in their work. Again, this may be connected to the nature of the work in NAV, which is often connected to vulnerable situations in the lives of citizens and is therefore regarded as more complex than taxation. The NAV data team has, therefore, initiated a project to standardize the overall process of impact assessment and juridical work. This data impact assessment contains a detailed description of the work processes and different phases of projects, in addition to risk descriptions tied to different variables. It also attempts to assess the potential societal consequences of the system as a whole.

The most interesting aspect is the perception by both data teams of gaps in the legal frameworks governing their work. This was also connected to the fact that, although there is a significant body of regulations dealing with privacy issues, machine learning regulation has often been left to vaguer ethical guidelines. These guidelines, however, are not the core domain of legal staff in public administration in these organizations. Making decisions concerning what are good enough reasons to use data, where data can be combined, and what decisions can be influenced by machine learning is left to individual organizations and data team members. They, however, are not comfortable making these decisions, while questions of fairness or transparency are not resolved by clear guidelines.

Data and infrastructure

Data both limit and enable the work of the data teams observed in this study. Policymakers and public sector leaders see the availability of data as a key enabler and force within the sociotechnical imaginary of public administration. According to these data teams, how data can be used depend on access, quality, and the context within which the data are collected. Both teams have expressed their requirement for resources to make data machine learnable. Data thus are mutually constituted and bound together in a variety of practices (Kitchin, 2014).

All policy and strategy documents examined for this study highlight the tremendous potential within the public sector's gold mine of data. However, existing materialities and infrastructures have not been constructed with this in mind. The data infrastructure of both organizations is, therefore, currently being upgraded. Although their efforts differ in size and original impetus, both public entities envision a new data platform to ease data analysis and the re-circulation and use of existing data. A visionary in the NAV IT department used the metaphor of a self-serve buffet to describe their project, where data could easily be picked and combined by a variety of actors. These infrastructures are envisioned to replace traditional data warehouses in both organizations. Both envision data platforms as strategic goals for the future. The NAV data team was required to build their own data infrastructure to start their work, to realize a sociotechnical imaginary on their own. As projects often do not start with data, but are required to be connected to a need or problem and a legal mandate in the organization, data enter the scene after a variety of apparatuses are already arranged. What data are available, and in what quality, is often unclear to the data teams. How to access previously unused data or acquire sufficient information on datasets is regarded by both data teams as difficult. Data are often not organized and curated for data analysis, in either organization. According to data team members, the data available in warehouses are often in aggregate form inadequate for machine learning purposes.

Mining public sector data is seldom a straightforward process. How much data can be put together depend on a variety of practical constraints. The sickness benefit project has spent nearly 2 years collecting and analyzing one data source: sick leave certificates. These have been produced by doctors around the country and submitted to NAV in various formats. In addition, hospitals have still not digitized all sick leave certificates. Although this project intended to use other data sources, it had to halt such efforts: "*This project is pragmatically driven. We do not have time to dig into more data right now. There are so many owners and sources*" (NAV data team member). Aligning data sources and cleaning the data so they can be analyzed requires significant work. The data team also pointed out that they have received a lot of feedback from case workers on the data sources they currently use in their assessment of cases. These sources, however, cannot be accessed by the team, as they lie outside of their systems.

The SKATT and NAV data teams spend a significant amount of time "getting to know" the data before starting their analysis. This activity is connected to understanding how data have been collected and why, and includes an active dialogue between subject matter experts and the data team. Discovering the limitations of the data is regarded as highly important by the data team members. This is clearly in direct contrast to the expectations of fast and cheap data analysis, and whole-population data, which feed the sociotechnical imaginary of datafied public services (Rieder and Simon, 2016).

Doing machine learning?

The data teams spend little time actual coding machine learning algorithms or training and testing models, as most of their time is spent determining organizational requirements, holding project meetings, acquiring data, managing data quality, performing legal

assessments, and discussing which model to use in both teams. "*Building the actual machine learning models, is what takes the least time here (laughs)*" (SKATT data team member). It is only at this point that the sociotechnical imaginary is finally realized into a specific working data assemblage, able to do work in the world (Kitchin and Lauriault, 2014). To date, few machine learning models have been deployed in these organizations. Most new data assemblages remain within the data teams and have not yet managed to do work in the world. At the time of observation, SKATT had been able to deploy seven models, while NAV still has none.

The actual training and testing of models has not been observed in this project. Nevertheless, the outcome and content of choosing, training, and testing was discussed with data team members in both public entities. Training and testing machine learning models is often experienced as highly interesting for team members: "*They love tinkering with the models*" (SKATT, team lead). The required accuracy level of machine learning models is decided on by the data team together with subject matter experts and legal staff. The sickness benefit project, for example, has decided that an 82% accuracy rate will be sufficient for the decision support tool. Again, a variety of subjectivities and communities is involved in these negotiations.

The day-to-day reality that a lot of time is not spent on any form of direct data work could be experienced as bothersome by practitioners. The NAV data team in particular expressed severe frustration at times. The content of the practice of machine learning is itself contested; several members of the data teams argued that all of the other work also is part of their responsibility and cannot be disregarded. This quote shows exemplifies the variety of tasks of a data team member:

> There were many processes to deal with and manual transfers of data and information. [. . .] The entire project consisted of many actors and it was our task to make this work. Client / management, tax assessment team, coordination team (management for the inspectors), project managers / analysts and controllers. In addition, one had to have an overview of the data flow in data warehouses and ask them to change their work routines. In other words, there was a lot of practical planning work. In addition, we were on tour to inform about the model and reassure the employees of what is going to happen. (SKATT, data team member)

Circulating and maintaining the sociotechnical imaginaries of data-driven public administration is regarded as equally important to actually building the data assemblage, at this point of time because the data teams depend on the acceptance of this future vision within their organizations. An analysis of presentations also shows how the data teams translate the dominant sociotechnical imaginaries developed by corporate and state actors into the context of public administration.

Interestingly, most data team members point out that machine learning or other predictive technologies may not be an appropriate solution to the problems they are assigned. The practice of machine learning, therefore, also includes the practice of understanding when machine learning is not the solution. In other words, data team members engage in negotiations concerning the extent and the boundaries of these sociotechnical imaginaries in their day-to-day work.

Discussion and conclusion: situating datafication

The datafication of public administration has been encouraged by both multinational and national policy and strategy documents for several years; the sociotechnical imaginaries related to this administrative reform have been circulated and been enacted in a variety of arenas. However, they largely represent visions of the future rather than any accomplished result. The context of Norwegian public administration prohibits a variety of datafication efforts. This situation is experienced as bothersome and frustrating by practitioners, who have set out to realize the promise of data-driven public administration. This observed friction, however, offers valuable insights into the variety of apparatuses that frame the nature, operation, and work of this administrative reform (Kitchin and Lauriault, 2014). Such insights contribute to our understanding of how data-driven technologies can be governed, controlled, and held to account (Dencik, Hintz, et al., 2019). A growing body of literature testifies to the societal consequences of datafication and shows how data-driven technologies are socially constructed rather than neutral (Beer, 2017; Kitchin, 2017; Redden et al., 2020). This article builds on the developing literature, reframing the apparatuses of the data assemblage by investigating how existing public administration apparatuses manage to constrain powerful future visions when actors attempt to translate and negotiate them into assemblages of materiality. This analysis demonstrates the rich observations made possible through situated, empirical studies on datafication in public administration in specific contexts, rather than attempting to make more abstract, general observations.

Sociotechnical imaginaries are contingent, multiple, and contested (Mager and Katzenbach, 2021). They often do not materialize in full, and not without interruption (Caplan et al., 2020). The data projects followed in this study are all intended to alter public administration but at the same time are themselves constrained by a variety of mediations. Structural and institutional constraints provide obstacles to the fulfillment of future visions of a data-driven society. While sociotechnical imaginaries clearly thrive in corporate discourse and policy, and within organizational discourse, the new data assemblages constructed by the data teams observed in this study remain highly unstable and demand immense resources (Hagendorff and Wezel, 2019). Practitioners are forced to constantly negotiate between existing apparatuses and broader future visions. The promise of a more effective and informed public administration is restricted by the technical, legal, organizational, methodological, and political constraints observed in these cases. Some of these constraints seem mundane at first but contribute to the mediation of sociotechnical imaginaries into small, often unstable, data assemblages. Datafication is always embedded into specific contexts, with structures, values, and conditions that are reflected in practice and thus already manage to mediate the power of data and algorithms.

The concept of the sociotechnical imaginary introduced in this research project offers an analytical tool to understand the nature and power of collectively held future visions, while the concept of the data assemblage directs our analytical attention toward the relational and contextual discursive and material practices shaping data-driven technologies. Currently, the constraints presented in this case study have prohibited the data teams from utilizing the vastly enhanced possibilities afforded by emerging technologies to understand, predict, and control the activities of citizens. At the same time, they also reduce the richness of

sociotechnical imaginaries, as can be observed in the narrow focus on control in projects. SKATT has a clear advantage when producing data-driven technology, as one of its main tasks is controlling, and public mandate is more clearly defined. Most constraints face contrary pressure from policymakers and other actors who work to maintain the sociotechnical imaginary and therefore aim to get rid of these obstacles. This has led, for example, to the idea of digitalization-friendly law, which is intended to enhance datafication (Plesner and Justesen, 2021). There is little reflection in the field on why these obstacles have been established in the first place. A deeper understanding of the interaction, interrelation, and negotiation between mediations is crucial to further develop the field of critical data and algorithm studies. To avoid oversimplified accounts of data-driven technology which are of little influence in this field of practice, research must take the variety of conditions of datafication into account (Moats and Seaver, 2019; O'Neil, 2016).

Data assemblages are not neutral and are influenced by a variety of apparatuses (Kitchin and Lauriault, 2014). Datafication may often be problematic, producing unintended consequences in its entanglement with power and politics (Beer, 2017; Bucher, 2018; Dencik and Kaun, 2020; Dencik, Hintz, et al., 2019). To understand datafication as more than a general, albeit substantial, shift in the organization of society requires us to situate it beyond Internet platforms. Data-driven technologies produced at the Google headquarters in Silicon Valley may differ substantially from monitoring systems in the Norwegian Tax administration. The critique of private sector datafication often seems to place society in passive or powerless positions, leaving little space for intervention and reification except after systems have already caused harm (Caplan et al., 2020). This study's approach avoids data-driven determinism, showing how policy, organizational structures, legal frameworks, subject matter experts, and existing data infrastructures are able to mediate datafication in significant ways. These constraints act as counterforces against dominant sociotechnical imaginaries strongly dominated by the private sector (Bareis and Katzenbach, 2021). The constraints laid out above provide us with an interesting point of departure to engage reflexively and critically in datafication practices. In many cases, data-driven public administration still remains a future vision rather than being realized in stable data assemblages; this might provide us, as citizens, users, or researchers, with opportunities to alter these imaginaries prior to or even during their translation into new and effective data assemblages. An actionable critique of the phenomenon might ask which limitations to interrogate and how these already existing constraints can and should be altered, and by whom. Further investigation will also require us to reflect more deeply on who is able to mediate them and who is not. The obstacles presented here can, therefore, act as points of entry to reconfigure the sociotechnical imaginary of data-driven public administration.

Author's note

The article is not currently being considered for publication by any other print or electronic journal.

Acknowledgements

The author thanks Heidrun Åm, Heather Broomfield, and Emil Røyrvik for reading and commenting on earlier drafts of the article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research presented in this paper was funded by the Norwegian University of Science and Technology's (NTNU) Digital Transformation initiative, project: 'Digital infrastructures and citizen empowerement (DICE).'

ORCID iD

Lisa Reutter https://orcid.org/0000-0002-4952-0416

Note

1. The research project had to constantly renegotiate access to the public entities and was heavily inspired by Nick Seaver's (2017) work on the ethnography of algorithmic systems. Here, Seaver argues to treat "access as a kind of texture, a resistance to knowledge that is omnipresent and not always the same." The topics of ethics and unintended social consequences have become quite prominent in the public sector: this was used to legitimate access to work practices in this study. While this article is theoretically driven and intended to contribute to the field of critical data and algorithm studies, other parts of the research project included more practically oriented research outputs and popular science presentations to practitioners.

References

Allhutter D, Cech F, Fischer F, et al. (2020) Algorithmic profiling of job seekers in Austria: how austerity politics are made effective. *Frontiers in Big Data* 3: 5.

Andrews L (2019) Public administration, public leadership and the construction of public value in the age of the algorithm and "big data." *Public Administration* 97(2): 296–310.

Bareis J and Katzenbach C (2021) Talking AI into being: the narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology & Human Values*: 1–27.

Beer D (2017) The social power of algorithms. *Information, Communication & Society* 20(1): 1–13.

Beraldo D and Milan S (2019) From data politics to the contentious politics of data. *Big Data & Society* 6(2): 1–11.

boyd D and Crawford K (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662–679.

Broomfield H and Reutter L (2019) *Kunstig Intelligens/data Science: En Kartlegging Av Status, Utfordringer Og Behov I Norsk Offentlig Sektor—Første Resultater*. Trondheim: Universitet i Oslo, NTNU, Difi. Available at: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2634733

Bucher T (2018) *If . . . Then : Algorithmic Power and Politics. Oxford Studies in Digital Politics*. New York: Oxford University Press.

Caplan R, Clark M and Partin W (2020) *Against Platform Determinism—A Critical Orientation*. Available at: https://points.datasociety.net/against-platform-determinism-899acdf88a3d

Dencik L (2019) Situating practices in datafication—from above and below. In: Stephansen HC and Trere E (eds) *Citizen Media and Practices*. London and New York: Routledge.

Dencik L and Kaun A (2020) Datafication and the welfare state. *Global Perspectives* 1(1): 12912.

Dencik L, Hintz A and Wahl-Jorgensen K (2019) *Digital Citizenship in a Datafied Society*. Cambridge: Polity Press.

Dencik L, Redden J, Hintz A, et al. (2019) The 'golden view': data-driven governance in the scoring society. *Internet Policy Review* 8(2): 1–24.

Elish MC and boyd D (2018) Situating methods in the magic of big data and AI. *Communication Monographs* 85(1): 57–80.

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Hagendorff T and Wezel K (2019) 15 challenges for AI: or what AI (currently) can't do. *AI & Society* 35: 355–365.

Hockenhull M and Cohn ML (2021) Hot air and corporate sociotechnical imaginaries: performing and translating digital futures in the Danish tech scene. *New Media & Society* 23(2): 302–321.

Jasanoff S and Kim S-H (2009) Containing the atom: sociotechnical imaginaries and nuclear power in the United States and South Korea. *Minerva* 47(2): 119.

Jasanoff S and Kim S-H (2015) *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. London: University of Chicago Press.

Kazansky B and Milan S (2021) "Bodies not templates": contesting dominant algorithmic imaginaries. *New Media & Society* 23(2): 363–381.

Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London: SAGE.

Kitchin R (2017) Thinking critically about and researching algorithms. *Information, Communication & Society* 20(1): 14–29.

Kitchin R and Lauriault T (2014) Towards critical data studies: charting and unpacking data assemblages and their work. The Programmable City Working Paper 2. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

Lehtiniemi T and Ruckenstein M (2019) The social imaginaries of data activism. *Big Data & Society* 6: 1–12.

Maciejewski M (2017) To do more, better, faster and more cheaply: using big data in public administration. *International Review of Administrative Sciences* 83(1_suppl): 120–135.

Mager A and Katzenbach C (2021) Future imaginaries in the making and governing of digital technology: multiple, contested, commodified. *New Media & Society* 23(2): 223–236.

Mayer- Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.

Mejias UA and Couldry N (2019) Datafication. *Internet Policy Review* 8(4): 1–10.

Mergel I, Rethemeyer RK and Isett K (2016) Big data in public affairs. *Public Administration Review* 76(6): 928–937.

Ministry of Local Government and Modernisation (2020) *Nasjonal Strategi for Kunstig Intelligens*. Available at: https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/

Misuraca g and van Noordt C (2020) *Artificial intelligence in public services—overview of the use and impact of AI in public services in the EU*. EUR 30255 EN. Luxembourg: Publications Office of the European Union.

Moats D and Seaver N (2019) "You social scientists love mind games": experimenting in the "divide" between data science and critical algorithm studies. *Big Data & Society* 6(1): 1–11.

O'Neil C (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing Group.

Pasquale F (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press.

Plesner U and Justesen L (2021) The double darkness of digitalization: shaping digital-ready legislation to reshape the conditions for public-sector digitalization. *Science, Technology, & Human Values* 47(1): 146–173.

Public Administration Act (2021) Act relating to procedure in cases concerning the public administration (LOV-1967-02-10). Available at: https://lovdata.no/dokument/NLE/lov/1967-02-10#KAPITTEL_3

Redden J (2018) Democratic governance in an age of datafication: lessons from mapping government discourses and practices. *Big Data & Society* 5(2): 1–13.

Redden J and Grant A (2020) Learning from cancelled systems. Available at: https://www.themj.co.uk/Learning-from-cancelled-systems/218519

Redden J, Dencik L and Warne H (2020) Datafied child welfare services: unpacking politics, economics and power. *Policy Studies* 41(5): 507–526.

Rieder G and Simon J (2016) Datatrust: or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society* 3(1): 1–6.

Ruppert E (2018) Sociotechnical imaginaries of different data futures. 3e Van Doornlezing. Available at: https://www.eur.nl/sites/corporate/files/2018-06/3e%20van%20doornlezing%20evelyn%20ruppert.pdf

Seaver N (2017) Algorithms as culture: some tactics for the ethnography of algorithmic systems. *Big Data & Society* 4(2): 1–12.

Skatteetaten (2025) (2017) Skatteetaten 2025: virksomhetsstrategi for strategiperioden 2018–2021. Available at: https://www.skatteetaten.no/globalassets/om-skatteetaten/om-oss/strategi/virksomhetsstrategi-2018-2021.pdf

Strauß S (2015) Datafication and the seductive power of uncertainty—a critical exploration of Big Data enthusiasm. *Information* 6(4): 836–847.

Tanweer A, Fiore-Gartland B and Aragon C (2016) Impediment to insight to innovation: understanding data assemblages through the breakdown—repair process. *Information, Communication & Society* 19(6): 736–752.

Tavory I and Timmermans S (2014) *Abductive Analysis – Theorizing Qualitative Research*. London: The University of Chicago Press.

Thomas SL, Nafus D and Sherman J (2018) Algorithms as fetish: faith and possibility in algorithmic work. *Big Data & Society* 5(1): 1–11.

Tupasela A, Snell K and Tarkkala H (2020) The Nordic data imaginary. *Big Data & Society* 7(1): 1–13.

van Dijck J (2014) Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. *Surveillance & Society* 12(2): 197–208.

van Ooijen C, Ubaldi B and Welby B (2019) *A data-driven public sector: enabling the strategic use of data for productive, inclusive and trustworthy governance*. OECD Working Papers on Public Governance. Paris: OECD Publishing. Available at: OECD Working Papers on Public Governance.

Williamson B (2018) Silicon startup schools: technocracy, algorithmic imaginaries and venture philanthropy in corporate education reform. *Critical Studies in Education* 59(2): 218–236.

Yeung K (2020) From new public management to new public analytics. ADM Workshop 1: Conceptualizing Automation in the Welfare Sector.

Ziewitz M (2015) Governing algorithms: myth, mess, and methods. *Science, Technology, & Human Values* 41(1): 3–16.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. London: Profile Books.

Author biography

Lisa Reutter is a PhD candidate at the Norwegian University of Science and Technology, researching the (re)construction of algorithmic governance in the Norwegian public sector. She is part of the interdisciplinary research project Digital Infrastructures and Citizen Empowerment (DICE). Her research is located at the intersection between the fields of public administration research, sociology, and science and technology studies.

[Paper 2]
Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation.

This paper is co-authored with Heather Broomfield and published in the *Scandinavian Journal of Public Administration* (2021) 25(2):73-97.

SJPA
25(2)

Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation

Heather Broomfield and Lisa Reutter*

Abstract

This paper aims to demystify the concept of data-driven public administration and lay bare the complexity involved in its implementation. It asks the overall research question of what challenges are encountered and problematised in a nascent phase of data-driven public administration implementation. The analysis is based on a multi-method research design, including a survey, follow-up interviews with practitioners and an analysis of key policy documents in the context of the Norwegian public sector. It highlights areas of both discrepancy and harmony between what has been prioritised at the policy level and the reality of implementation on the ground. In addition, unseen issues are discussed in order to broaden this perspective. Data-driven administrative reform touches upon everything from organisational culture to technical infrastructure and legal and regulatory frameworks. The complexity laid out in the analysis thus has implications for theory and practice. Nordic countries provide an interesting object of investigation, as they hold vast amounts of data and are highly digitalised, yet, in common with many other governments, they are still in a nascent phase of implementation. This paper should therefore be relevant to other jurisdictions and it provides a call to arms for civil servants and public administration scholars to engage more deeply in this phenomenon.

Heather Broomfield,
Department of Public and International Law, University of Oslo, Norway,
heather.broomfield@digdir.no

Lisa Reutter,
Department of Sociology and Political Science, Norwegian University of Science and Technology, Norway,
lisa.m.reutter@ntnu.no

Keywords:
public administration reform,
data-driven,
artificial intelligence,
policy implementation,
unseen issues

Introduction

The ambiguous, multifaceted and contested nature of data-driven public administration presents a serious challenge to practitioners, policymakers and scholars alike, ushering in a new and all-encompassing chapter in the extensive history of public administration reform (Bullock 2019). This paper aims to demystify the concept and provides a unique account of the Norwegian public sector's early endeavours to implement data-driven government. The paper asks the overall research question of what challenges are encountered and problematised in a nascent phase of data-driven public administration implementation. It identifies and discusses these challenges as experienced by practitioners. We highlight areas of both discrepancy and harmony between what has been prioritised at the policy level and the reality of implementation on the ground. The paper then goes on to discuss issues that are largely unseen by both policymakers and practitioners, but that scholarship has identified as potential unintended consequences caused by the utilisation of data-driven technology in the sector. We aim to both lay bare the complexity involved in implementation and convey the paradigm shift that this is bringing to public administration. This paper provides valuable insights for civil servants and policymakers embarking upon their own data-driven journeys. It also issues a

***Heather Broomfield** is a PhD candidate at the University of Oslo, Department of Public and International law. She is also a senior advisor at the Norwegian Digitalisation Agency. Heather is researching the framing and governance of a datafied public sector.

Lisa Reutter is a PhD candidate at the Norwegian University of Science and Technology, department of sociology and political science, researching the datafication of the public administration. Lisa is associated with the interdisciplinary research project Digital Infrastructures and Citizen Empowerment (DICE). Her research is located at the intersection between the fields of public administration research, sociology and science and technology studies.

Scandinavian Journal of Public Administration
25(2): 73 - 97

ISSN: 2001-7405
e-ISSN: 2001-7413

call to arms for public administration scholars to engage more deeply with this phenomenon.

The amount, granularity, immediacy and variety of data about subjects to be governed is unique for modern governments (Ruppert, Isin and Bigo 2017), and are resources in which the Nordic public sectors are deemed to be particularly rich. There is a palpable sense of urgency in Norway to utilise this data goldmine. It is prescribed as a treatment to alleviate the consequences of impending threats to the Norwegian welfare model caused by issues such as demographics, downswings in the oil sector and increasing immigration rates (Dølvik et al. 2014). Data-driven public administration aims to promote the idea of data as an asset that needs to be highly integrated into policy-making, service delivery, organisational management and innovation (van Ooijen, Ubaldi and Welby 2019). It carries with it the promise of effectiveness and improved services. As with many other governments, Norway is in the nascent phase of this transformation and is far from unique in its quest, as this idea is also highly advanced by the OECD and other transnational actors (Misuraca and van Noordt 2020; Sun and Medaglia 2019).

Industry and governments are enthusiastically embracing technological trends such as platforms, the cloud and machine learning, seeking to harness what is perceived to be a tremendous potential in technology and data (Yeung 2018). It appears that new technology and technology-related practices are sweeping over public administration without being critically assessed by the field. A significant amount of research has been conducted on the issue of becoming data driven, but this research is predominantly concerned with private sector actors and an outside perspective on the phenomenon. Where public administration research is emerging, it is dominated by UK and North American case studies. Scandinavian governments, however, operate in a different data context. They collect and manage vast amounts of detailed personal and nonpersonal data and experience relatively high levels of trust from citizens. This specific context requires empirical and theoretical attention. Information technology has long been neglected in public administration research, leading to the marginalisation of the discipline's influence on practical policy-making (Dunleavy et al. 2005). Despite the increasing dominance that both digitalisation and data-driven approaches have over public administration and in public policy, there is little to signify that the situation has improved much in recent years, with many researchers lamenting the relative dearth of research concerning this "new" era of data-driven government and expressing the urgent need to engage (Agarwal 2018; Brauneis and Goodman 2018; Redden 2018; Wirtz, Weyerer and Geyer 2019). This lack of research may also be contributing to the discursive context of data-driven public administration being significantly shaped by corporate technology companies (Andrews 2019).

This paper adopts a practice approach, focusing on the "perceived challenges to act upon" as identified by public service practitioners tasked with realising the promised riches from the "goldmine" of data. It focuses on the framing of this problem rather than its resolution (Andrews 2019). The data material consists of a survey, follow-up interviews with practitioners and analysis of key policy documents. As advocated by Veale, Van Kleek and Binns (2018), we conduct this research in collaboration with those at the coalface,

rather than working from afar, in order to endeavour to elicit new insights and understand aspects that may not be immediately apparent from the outside. We discuss data-driven public administration as a complex practice that challenges traditional public administration and therefore impacts upon everything from organisational culture to technical and data infrastructure and legal and regulatory frameworks.

The first part of the paper introduces and discusses the concept of data-driven public administration; it then proceeds to situate the phenomenon in implementation research in addition to providing a summary of the Norwegian data context. We then advance our mixed-method approach. The analysis provides an overview of the current state of Norwegian data-driven public administration and the challenges encountered in this nascent phase. As the invisible labour and ambiguity behind the grand ideas of data-driven government are laid bare, this may contribute to a more balanced and practice-based approach to the phenomenon.

What is data-driven public administration?

The realisation of the optimal modern, responsive, efficient public administration deemed to have evaded us thus far, is envisaged to be delivered by "data-driven government". A term popularised by the corporate sector and transnational actors and subsequently adopted by many governments. It is defined as follows by the OECD.

> *"A data-driven public sector recognises data as a strategic asset in policies and services design and delivery. It implies the development of sound data governance structures (including data strategies, institutional arrangements, rules) and related delivery mechanisms (data infrastructures, standards) to capitalise on the value of data to anticipate and respond to the needs of users, deliver better services and policies, and promote data integration, access, sharing and use across the public sector. A data-driven public sector also favours the use of innovative and alternative sources of data in the evaluation and monitoring of policies and services over time." (Ubaldi et al. 2020: 30)*

The modern state and data are already inevitably woven together (Desrosières 1998). New public management has, in addition, increased the focus on quantification in the sector (Muid 1994). Data is generated, managed, stored, processed and analysed in every aspect of public administration, the state being a key producer, provider and consumer of data (Kitchin 2014). Reform, redesign, reinvention and a perceived urgent need to adjust to rapidly changing circumstances in the sector are also not new, public sector innovation having become a professional default state (Wagenaar and Wood 2018). Information technology is now recognised as a key instrument of administrative reform (Kramer and King 2006; Margetts 2009). What, then, is new with this idea of a data-driven public sector?

The idea of data-driven public administration builds upon but goes far beyond current ubiquitous processes, such as digitalisation, e-government and evidence-based policy-making. It consists of two interwoven processes: the use of more and different data and more advanced methods to analyse this data (artificial intelligence [AI], machine learning, etc.) to feed it back into existing work processes. The calculative systems and techniques to process information have become ever faster, more comprehensive and more autonomous in recent years (Beer 2017). ICT once utilised for data entry are now capable of cognitive and analytical tasks, bringing with them the potential to automate many aspects of public sector work, such as policing, nursing and teaching (Bullock 2019; Busch and Henriksen 2018). They are moving from rule-based systems to finding patterns in data, allowing for automated decision-making and providing decision support tools. Knowing, data-driven and predictive technologies are part of a shift towards automated, anticipatory, and algorithmic forms of governance (Williamson 2014). This leads to profound changes to the way in which the public administration learns about, engages with and responds to citizens (Redden 2018; Hintz, Dencik, Wahl-Jorgensen 2019).

Data-driven public administration, nationally and internationally, promises endless opportunities, the rhetoric reflecting what Elish and Boyd (2018) describe as boundless, seasoned with a sort of magic. It provides the sector with a sense of being able to do more, better, faster and more cheaply through automation or augmentation and is perceived as a solution to respond to the growing complexity of society (Klievink et al. 2017; Maciejewski 2016). Examples of use areas include more personalised and context-based welfare (to ensure that state benefits go to the most vulnerable families) and autonomous vehicles to revolutionise public transport. Disease can be more accurately diagnosed and treated; control and fraud detection can be greatly improved; children most likely to be at risk from abuse can be identified and followed up upon; faster and richer images of evolving reality can be provided, allowing for natural disasters to be better predicted and managed; terrorist attacks can be prevented and traffic congestion relieved (Barth and Arnold 1999; Bullock 2019; Klievink et al. 2017; van Ooijen, Ubaldi and Welby 2019).

However, experience to date with data-driven technology in the public sector is peppered with examples in which this technology has had significant unforeseen and unwelcome consequences. Some examples here are errors in cancer screening in the UK (Andrews 2019). An automated system for detecting welfare fraud was found to violate human rights in the Netherlands (Henley and Booth 2020). Students taking the International Baccalaureate (Schei 2020) and the Leaving Certificate in Ireland (Lillington 2020) were assigned incorrect grades. Predictive policing in the US resulted in the racial targeting of black neighbourhoods due to biased data, and teachers have been unfairly dismissed and their competence undervalued due to algorithm scoring in schools (O'Neil 2016).

A number of negative, albeit often unintended, consequences have thus already been pointed out by research. These include impenetrable opaqueness, reinforcement of discrimination and the facilitation of surveillance (Alston 2019; boyd and Crawford 2012; Kitchin 2014; Pasquale 2015; van Dijk 2014). Concerns are raised that it is causing a change of power dynamics between state

and citizen, as data-driven public administration increases the ability to understand, predict and control citizens' behaviour (Hintz, Dencik, Wahl-Jorgensen 2019). These issues have, however, received little attention from policymakers, practitioners and public administration scholars alike.

Implementing data-driven public administration

Data-driven public administration is envisaged as an all-encompassing public administration reform, fundamentally changing the way democratic systems engage with and learn about citizens (Redden 2018). It is expected to be integrated into all aspects of public administration, from policy-making to service delivery, from organisational management to innovation (van Ooijen, Ubaldi and Welby 2019). It is not one concrete policy from which to neatly analyse but instead a more general latent trend of data-driven public administration. The increased influence of the imaginary of data-driven public administration must therefore be situated within the wider field of public administration research. Reformers are often over-optimistic, hold unrealistic expectations and fall into the many traps of implementation (Caiden 1999). Data-driven public administration is in the nascent phase of implementation, and it is therefore too early to assess or evaluate this reform; instead, we endeavour to illuminate and explain but not to affect what happens (Caiden and Puniha 2011; Hill and Hupe 2014). We therefore situate this paper within implementation research to expose "what happens between the establishment of policy and its impact in the world of action?" (O'Toole 2000: 266). This paper thus exposes the challenges and reveals the complexity involved by studying how the policy expectations of data-driven public administration are being implemented by practitioners on the ground, thereby embracing both a top-down and bottom-up perspective as advocated by De Leon and De Leon (2002).

The analytical framework focuses on current implementation challenges. We firstly identify the experience of practitioners and how policy is being put into action. We then situate these within the policy context, in order to identify areas of discrepancy and harmony between what has been identified and prioritised at the policy level and what is actually happening. As De Leon (1999: 322) points out, "The main problem with implementation is that the discrepancy between 'something' and 'that idealized thing' is often a matter of rose-coloured expectations". We then proceed to discuss some of the unseen areas that research has pointed to but that neither policymakers nor practitioners have prioritised.

A number of practical challenges in the implementation of data-driven public administration have already been identified by scholars. These involve issues at the system, organisational and individual levels (Pencheva, Esteve and Mikhaylov 2018). The challenges are therefore not limited to technical issues but also include ethics, processes, analytics and organisational and institutional change (Mergel, Rethemeyer and Isett 2016). Issues addressed in earlier research include uncertainty tied to fairness, accountability and discretion (Veale, Van Kleek and Binns, 2018); unrealistic expectations towards AI and a lack of interdisciplinary talent (Sun and Medaglia 2019); hidden costs produced by data-driven technology (Hagendorf and Wenzel 2019); and unanswered questions around ethics and democratic governance (Mergel, Rethemeyer and Isett 2016).

Fredriksson et al.'s (2017) literature review on big data in the public sector identified three main challenges in data application, namely, the management of data, ensuring data quality and ethical and privacy concerns tied to the use and sharing of data. Redden's (2018) case study on the Canadian public sector employed a counter-mapping approach consisting of freedom of information requests, semi-structured interviews and document analysis to map issues raised by public administration in its work towards data-driven public administration. She identifies a variety of practical challenges, such as the technical infrastructure, access to data, privacy and security, skills gaps, organisational culture and data quality and accuracy. These challenges provided a foundation for the mapping of the Norwegian public sector. Practitioners and policymakers are often unaware of the full range of practices and related challenges (Wirtz, Weyerer and Geyer 2019). Studying this early phase of implementation provides many entry points for further research and a basis for public administration scholars, who are currently underrepresented, to engage in this field. As Barth and Arnold (1999: 349) find,

> *"the real danger of [data-driven technology] in government is represented by researchers who are divorced from the world of public administration scholars and practitioners who are engaged in discussions and making technological decisions, without understanding the implications for governance of the administrative state."*

Civil servants and policymakers who are embarking on the data-driven journey can also gain valuable insights for their context and can be inspired to look more deeply into how the problem may be framed going forward. As implementation is highly context dependent, it is, however, important to point out some distinctive characteristics of the Norwegian data context in order to provide a foundation to assess the relevance of our findings for other jurisdictions.

The Norwegian public sector data context

Recognition of the value of data as a resource for the entire public sector is not new in Norway or in the Nordics. The sector is characterised by a long tradition of systematic data collection. At a time when many other jurisdictions shied away from data collection, Norway and the Nordic region embraced it, deeming it necessary for the establishment and good functioning of the welfare state. A personal identification code unique identifier system, established in 1961, assigns everyone a number either at birth or upon immigration. Originally incorporated into the National Population Register, it is now used extensively in hundreds of registers, such as those of education, employment, health, tax, social welfare and crime to name but a few. These registers operate under a legal mandate, with organisations assigned the responsibility of managing them, resulting in a high level of quality (Tupasela, Snell and Tarkkala 2020). With few exceptions, such as the common contact register, citizens are not permitted to opt out, given the registers' central role in the functioning of the state. It is therefore impossible to avoid leaving traces in administrative registers (Hovde Lyngstad and Skardhamar 2011), thereby allowing for a continually growing

data archive of the entire population. The use of data from these personal registers is currently strictly regulated, and the data is siloed.

Data collection extends beyond individuals to many other aspects of the economy and society, such as an extensive company register, a road and traffic database and the ordnance survey. National archive data stretching back for centuries is currently being digitalised, and, due to many years of digital public service delivery, there are vast amounts of behavioural data stored by the public sector, which is commonly termed "data exhaust".

The Norwegian public sector is particularly sectoral and somewhat fragmented. It comprises 70 executive agencies, 16 national ministries and 358 municipalities. There is a high degree of autonomy, with strict boundaries at the organisational, sectoral and municipal levels, and decision-making is largely consensual. It is also highly digitised and experiences high levels of trust from citizens (OECD 2017).

There have been a number of initiatives to improve data coordination, evidenced by many documents stretching as far back as 1988 (NOU 1988:40). These had limited tangible success. There is now, however, a palpable sense of urgency amongst practitioners, politicians and policymakers alike, who are salivating at the prospect of breaking down the data siloes and tapping into this treasure trove of data for secondary uses.

Methodology

This study employs a practice approach, focusing not only on the practitioners engaged on a quest to produce data-driven practices but also on the institutions and policies guiding these approaches (Dencik 2019). To be able to answer the overall research question of what kind of challenges are encountered and problematised by practitioners and policy, we advocate for a multi-method approach. Data-driven public administration is still a contested concept within public administration and, therefore, challenging to operationalise. In order to be able to obtain concrete examples of and challenges to data-driven administrative reform, we chose to tie this to the development and implementation of AI and data science in a survey and interviews, as these concepts are well known to our informants.

This study was conducted in 2019, a time when interest in AI adoption and data-driven public administration by the Norwegian public sector blossomed, catalysed by the initiation of work on a national AI strategy. The initial aim was to obtain a general understanding of how diverse entities engage in new data practices and the challenges that practitioners meet. This resulted in a practical report on the challenges of AI and data science, which is available online[1]. The empirical account given in this paper is thus based on an in-depth secondary analysis of the data material, which consists of a survey answered by practitioners (n=35) in 26 public organisations, follow-up interviews with 12 of the entities and a document analysis of relevant policy documentation (See Appendices A and B for the survey template and interview guide.) It is thus a two-phase explanatory research design (Creswell and Clark 2011). The practitioners who provided input to this study are system-level designers, not street-level bureaucrats, as most data-driven efforts can be observed at this level.

The survey aimed to map the status in the sector and provide the project with descriptive statistics, asking what projects were being initiated, what the data-driven practices were intended to be used for and which challenges the practitioners perceived as important to act upon. The participants were recruited from the Norwegian public sector AI forum, a meeting place for practitioners engaged in data-driven practices. This forum had 46 member organisations at the time, all of which were invited to participate. The response rate was 56%. The forum comprises agencies that are either planning data-driven practices or have already deployed them. The sample, therefore, consciously consists of already quite digitally advanced entities within the Norwegian public sector. This, of course, could indicate that the organisations may have already overcome some obstacles. Seventy-three percent of the responding organisations had fewer than 500 employees in total. The sample is small in size, unfit for quantitative analysis beyond a descriptive overview. Several practitioners reported that they responded to the survey in groups of two to three in order to discuss their responses with colleagues. The participants were guaranteed anonymity to encourage openness. Leadership was not included. The main question of the survey consisted of 13 general challenges when working with data-driven public administration; these were inspired by challenges identified in earlier research. The practitioners were asked to rate each challenge on a Likert scale from 1, "no challenge at all", to 5, "a very big challenge". The study was not limited to these, as 73% of the respondents either elaborated upon or reported other challenges in the free text field provided. The survey answers varied significantly and provided a unique composition of experiences and challenges for each of the public entities.

In the survey, invitations were offered for a follow-up interview. Forty-six percent of the organisations accepted; 33.3% of these had more than 500 employees and can therefore be categorised as large public entities in the Norwegian context. The interviews served to contextualise and advance our understanding of the practitioners' experiences and were conducted approximately three months after the survey. The interviews were conducted in a semi-structured manner with either individuals or small groups. How many participated in the actual interview was left to their discretion. The interviews lasted between 60 and 90 minutes. They were recorded, transcribed and anonymised.

This project has not interviewed or surveyed policymakers. Instead, the public sector digitalisation strategy (Kommunal-og moderniseringsdepartementet 2019) and the concept phase analysis (Difi 2018) were analysed as key policy documents, as these are regarded as the main guiding documents by practitioners due to their relevance for the public sector and their topicality. We identified issues as "Highly prioritised by policymakers" both by looking at the number of times these issues were mentioned in policy documents and by considering the emphasis that was placed on the issues, both discursively and through asking for concrete actions in the policy documents.

Whilst the survey provided the project with a general overview of efforts and challenges, the follow-up interviews further contextualised these and linked these challenges within the practitioner's discourse and experience. The

interviews were analysed inductively according to meaning, and the interviewees were thus treated as respondents. The policy analysis then allowed us to identify how the expectations at the policy level were being implemented on the ground by focusing on the conceptualisation and prioritisation in our discourse analysis of the documents. Combining the practitioner and policy perspectives also provided us with an overview of unseen issues at both levels. Together, the survey, interviews and policy provided this research project with a unique insight into the inner workings of the implementation of data-driven public administration at the system, organisational and individual levels (Pencheva, Esteve and Mikhaylov 2018).

Towards a data-driven public sector

The results of the analysis show that the approaches, status and perceived challenges to the realisation of data-driven public administration differ significantly amongst practitioners, reflecting the fragmented nature of public administration more generally.

Data-driven public administration is regarded as central across the sector, capable of transforming all aspects of public administration. It is regarded as the inevitable future. As one interviewee stated,

> *"We think about automation, we think about efficiency, we think about getting rid of all routine tasks, which we do not need to do. Let's get a machine to do it." (Interview 7, small public entity)*

The sector is, however, very much in the nascent phase. There are few data-driven technologies in production, with most still in the planning or piloting phase. Just four out of the 26 organisations had systems in production. Control and risk assessment in case work, the automation of routine work and the potential for new predictive services are the predominant use areas being considered. Opportunities for predicting citizen and organisational needs were particularly highlighted. Decision-support tools and fraud control are the most common areas of application. Most of the organisations are working towards the integration of data into service delivery and organisational management rather than towards its contributing to policy planning in this early phase.

The participants regarded the main aim of a data-driven public administration to make the sector more effective, which translated to resource allocation and an improved response to user needs. There appears to be a consensus that most aspects of public administration can be standardised and data-driven and that data-driven public administration is not a substitute for but, instead, a supplement to traditional case work. Most do not consider AI and data science as goals in themselves but rather as necessary tools for realising the bigger idea of data-driven public administration. Several participants seem highly influenced by the private sector, referencing conferences and industry reports as their source of inspiration.

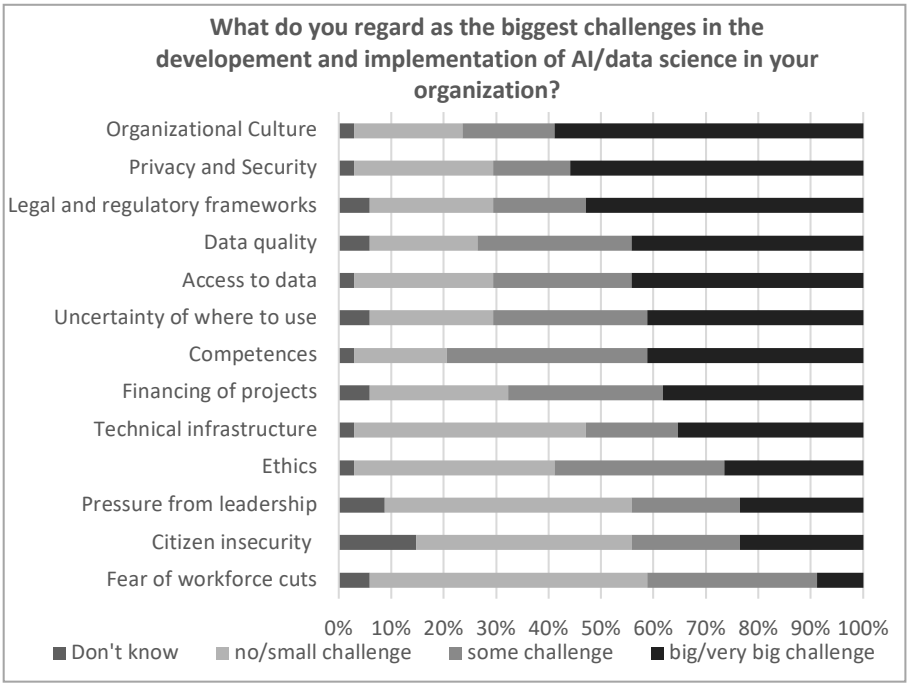
There is concern that there is a lack of appreciation at both the policy and organisational level around the level of investment necessary to realise data-driven public administration. Many struggle to sufficiently finance their efforts.

Initiatives change work processes; resources are therefore necessary to incorporate these into the organisation. The public sector is predominantly project oriented, but data-driven public administration is a process that does not fit neatly into short-term project thinking. This is a common challenge in policy implementation in the Norwegian context and far from unique to data-driven public administration (Dille and Söderlund 2013). Often, small projects are initiated and evaluated but halted when they need scaling up, which is a contributory factor in many projects' continuing to be in a pilot phase. Despite the excitement at the policy and leadership level, the reality is that the goals of data-driven projects are often unclear and unpredictable and therefore incompatible with current performance management regimes. Many struggle to understand how they can and should apply data-driven public administration and lack a clear understanding of what it actually means in practice, questioning whether their organisation is actually mature enough to adopt it. Aligning reality to expectations is difficult.

Perceived challenges to act upon

Data-driven practices introduce both challenges and opportunities to public administration. An overview of perceived implementation challenges, ranging from the most to least important as perceived by practitioners in their survey responses, can be found in Figure 1. The primary concern was organisational culture, followed closely by privacy and security and regulatory challenges. The interviews revealed that each of the identified challenges includes a subset of issues and that interpretation varies widely. In addition, the perceived importance differs across entities.

Figure 1. Mapping of perceived challenges to act upon



This section groups and discusses the perceived implementation challenges and prioritisations at both the practitioner and policy levels. We examine where practitioners and policymakers are aligned, where they diverge and where issues are rarely mentioned in the discourse. We begin with "Infrastructure, Access and Quality", which are highly prioritised at the policy level but less so at the practitioner level. We then take "Law, Privacy and Security", where the parties are aligned, before proceeding to discuss "Organisation, Internal Culture and Competence", which practitioners are particularly concerned with. Finally, we look at what we term the "unseen": issues that are rarely discussed but are of concern to researchers in the field.

Infrastructure, access, and quality: the policy darlings

National policy is increasingly focused on the potential for a data-driven public sector. The sharing of public sector data and investment in infrastructure are deemed central to the achievement of this goal, the expectation being that releasing this "raw material" will automatically realise transformation. It will usher in a more efficient administration, make citizens' and businesses' lives easier and simultaneously increase value creation in the private sector (Kommunal-og moderniseringsdepartementet 2019). Current actions and plans profess that a legal and technical infrastructure facilitating the sharing of high-quality data will transform Norway into a data-driven leader.

Data-driven public administration is predominantly perceived by policymakers as a technical issue, requiring technical infrastructure to provide access to high volumes of good quality data. This manifests itself in the prioritisation of the material aspect of the task, such as the purchasing of cloud solutions and building national data and API catalogues and data lakes. Our findings show that access to data and technical infrastructure are, however, considered less important by practitioners than the current policy assertions and general discourse would lead one to expect. It would be remiss to assert that it is not an issue, as instead it is one of a myriad of challenges that organisations face. Practitioners often actually uttered a sigh of "enough data already", as it is not necessarily the technical solutions and availability of data that are hindering their progress. The sheer existence of data and an infrastructure to access it thus does not magically enable a transformation of practices.

Data quality is also highly prioritised at the policy level, based on the assertion that bad data produces bad results. As is the case with data access, quality is an issue amongst practitioners, but it does not reflect the dominant position that it enjoys at the policy level. Again, this is not to assert that it is unimportant but instead that, when using data for data science purposes, many other factors also need to be considered that go beyond quality. These issues include but are not limited to contextualisation, data bias, suitability for secondary purposes and downstream issues (Veale, Van Kleek and Binns 2018). Data can be technically correct with high quality yet still be problematic. These challenges are complex and interdependent. The interviewees were equally concerned with these issues as with the actual quality of the data.

The emphasis on access, infrastructure and quality being the fundamental requirements to realise the holy grail of data-driven government seems to be oversimplified and reflects a deterministic view of the issue. Our findings show

that there is a discrepancy between policymakers and practitioners here. Data-driven public administration is far more than a technical issue. As the complexity of the task is discussed in the following sections, it becomes apparent that many of the other issues are less tangible, less measurable and cannot be as easily communicated as investment in a technical infrastructure, in which the number of data sets shared can be counted and individual organisations' "progress" can be measured.

Law, privacy and security: the bothersome -where policy and practitioners align

Practitioners and policymakers alike are particularly concerned with legal and privacy and security issues. Both scored high in survey responses and were much discussed in interviews. Data-driven practices are considered to be inhibited by current law; many respondents indicated that getting permission to access much of the register data for secondary purposes is particularly challenging. The current legal regime is perceived to be outdated and not fit for purpose to realise data-driven reform. Several entities are taking action to create more general legal mandates to allow for greater public sector data sharing and use. One interviewee even suggested that the prospect of establishing the public sector as one entity under GDPR, to enable free sharing of personal data, is being openly discussed within the public sector legal community. There is the utmost respect given to the notion that access must be balanced to protect privacy and security. Many struggle to design adequate privacy impact assessments for their work and argue that there are insufficient guidelines that consider how data should flow within and between organisations. They request assistance on issues such as the anonymisation and synthesisation of data, and many mentioned the need for regulatory safe spaces to experiment and gain experience with advanced technologies, highlighting that such "regulatory sandboxes" could form the basis for intersectoral cooperation.

Legal issues were also set within the interpretation context, framed around the individual behaviour of internal lawyers. Those taking a broad interpretation that allows for leeway in the law for the use of data were predominantly considered as progressive, and those who were stricter were described as conservative and as hampering progress. This distinction was particularly apparent and frustrating for the less-experienced organisations. The more experienced seem to have a mutual respect and cooperation with their lawyers. Regardless, the absence of common, streamlined interpretations of the law contributes to unpredictability across the sector. The general lack of competence amongst lawyers to understand the technical capabilities of how data-driven practices work was also identified as a weakness by all.

These findings echo those at the policy level. The digitalisation strategy calls for clear and digitalisation-friendly regulation and for a resource support centre to increase the sharing of data (Kommunal- og moderniseringsdepartementet 2019). This alignment is driven primarily from the perspective of data sharing and the protection of personal data. However, many other fundamental issues relevant to regulation and many classic dilemmas for the public sector come into play when embarking on data-driven practices. Examples include the public sector definition of fairness and explainability, bias, transparency,

accountability, discretion and broader challenges in the safeguarding of basic values in the Norwegian model, such as universality and the protection of vulnerable groups. Whilst these are not completely missing from the discourse, it is fair to say that a deep consideration of these issues in the legal and regulatory context is lacking.

Organisation, internal culture and competence: practitioner obsessions

Organisation, internal culture and competence are major issues for practitioners. Organisational culture itself encompassed a variety of ideas, with many respondents deeming that the realisation of data-driven public administration is hampered by internal resistance to change. Many references were made to age profiles, with older members of staff considered reluctant to embrace data-driven public administration and resistance attributed to a traditional mindset amongst both domain experts and leadership. One interviewee described it as follows.

> *"Organisational culture is still a challenge. I see it as a huge challenge. That's because we have less time than natural retirement will help us with, so we have to make changes. In fact, we must initiate great change in the entire organisation." (Interview 5, small public entity)*

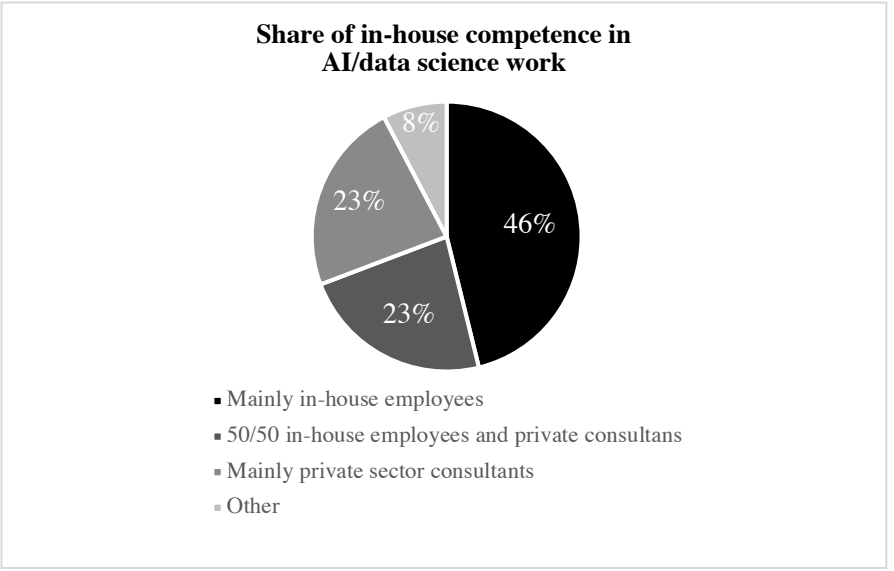
This was also observed in the Canadian case, where Redden (2018) expressed concern that the designation of internal reluctance as a "culture clash" is lamentable, as it may silence legitimate concerns. What also might be missed here is that it is easier to assign blame for slow progress on reluctant bureaucrats than to acknowledge the complexity of the process.

Each organisation asserted the necessity for multi-disciplinary cooperation, with all noting the significance of involving domain knowledge, in addition to data science and IT competence, from the outset. However, there is a spectrum here. The more experienced the organisation, the greater the emphasis on involving domain experts seemed to be. Bringing this multi-disciplinary cooperation from an assertion to a reality is fraught with difficulty. Many attributed this to different perspectives, the absence of a common understanding and lack of internal experience with multi-disciplinary cooperation. A further observation is that "multi-disciplinary" translated to involving technical, legal and domain knowledge competencies. Few advocated for the need for public administration and social science competencies, which Mergel, Rethemeyer and Isett (2016) point out is necessary, given their substantive depth on research methods and theory and the understanding of potential unintended consequences.

Many practitioners, whilst agreeing on the importance of data science competence, also expressed concern regarding the need for data competence within leadership, legal and domain knowledge experts. This echoes Veale, Van Kleek and Binns's (2018) observation that a lack of knowledge amongst those vertically accountable for service delivery hampers progress. Many expressed the view that ownership of and responsibility for the identification of opportunities should lie with the business side; however, the lack of competence on the potential of the technology is a hindrance. "No one really sees a need for this" was mentioned by several participants. A considerable amount of time is therefore spent in some of the more advanced organisations to increase the

technical expertise of domain knowledge professionals, and this is viewed as a critical success factor.

Figure 2. Competence profiles in AI/data science production and implementation



A key concern in much of the current research is fear of the influence that the private sector will have on public services, due in part to a lack of competence in the sector, a perceived inability to attract in-house competence in a pressed employment market and the political prioritisation to outsource (boyd and Crawford 2012; Brauneis and Goodman 2018; Redden 2018). The situation on the ground, however, is not as clear cut as current research suggests. Survey responses show that only 30% of the entities mainly use private sector consultants (see Figure 2). All the interviewees pointed to the importance of in-house expertise in the development, ordering and implementation of data-driven practices. Despite political prioritisation to outsource, experienced practitioners are reluctant to use private consultants for the entire process. The justification here was threefold. Firstly, they see a lack of understanding of the public sector and the type of data and responsibility involved. Secondly, data science is a long-term issue requiring regular development and maintenance. Thirdly the public sector needs to understand and control what is being done and cannot simply outsource this. A hybrid solution, if possible, seems to be preferred. The general assumption that the public sector finds it difficult to attract data science expertise is also nuanced. Whilst there is some truth in this, it was not experienced across the board. For many technologists, the public sector is an attractive "workplace with meaning", and many of our informants (and particularly those from the larger and more advanced organisations) are inundated with applicants to positions. Smaller organisations, by contrast, are struggling to recruit. This suggests that they may be forced to utilise the private sector for the entire data-driven process and risk that the private sector may shape public sector actions.

In addition, as Andrews (2019) points out, data-driven practices are often framed around leadership attitudes. Our discussions identified three distinct groupings of leader perspectives: some experience prioritisation difficulties, as leaders fail to see value; others encounter unrealistically high expectations; in the third grouping, leaders see recent data-driven capabilities as simply a new form of digital technology to be incorporated into the organisation rather than as a fundamental change in direction needing specific leadership prioritisation—an evolution rather than a revolution. This latter category was particularly apparent in the organisations that have considered themselves data driven for many years. Another observation is a lack of clear distribution of roles, responsibilities and authority. Many see a need to restructure their organisation, with the responsible unit often randomly placed. It can, for example, be found in the IT, statistics or analysis department.

All expressed the importance of the public sector AI forum, which was established by practitioners and which specifically does not permit membership from the private sector. This is an area for exchange of competence and is seen as a safe and open environment in which to learn from both success and failure.

Discussion: The unseen

The analysis identified many challenges encountered in the implementation of data-driven public administration. The policy level places significant emphasis on data and infrastructures, whereas practitioners are more concerned with organisational and competence issues. Both levels regard legal frameworks as a major hindrance. Discussing unseen issues in government discourse is important in order to broaden the perspective (Redden 2018). We observed that several of the potential negative implications and risks identified by scholarship were rarely considered at either level. These challenges are often more value laden and less practical, and procedural, making them more difficult to grasp for public sector actors. Nevertheless, highlighting the unseen and connecting it to the practical issues highlighted above adds to a procedural and critical understanding.

The potential for changing power dynamics between citizen and state as well as the insecurity over citizens' growing concern regarding the use of data-driven practices—as risks that could weaken trust (Redden 2018) - were rarely addressed. There is consensus at all levels that work be conducted in such a way that the high levels of trust that the sector enjoys should not be diluted. However, trust was often equated to privacy and legal issues. The perception seems to be that, when privacy is protected and current regulations adhered to, there is little concern. The practitioners were well intentioned, with many stating that they were "using technology for the good of society", but this is a normative assessment without concrete content and guidelines, and effectiveness was clearly a dominant variable. A data-driven meeting with the public sector requires large data flows between public entities in which the citizen becomes more visible, even if data is managed and shared responsibly (Hintz et al. 2018). One interviewee was particularly concerned with how far the public sector should go here, stating, "Society needs to agree with itself about what it wants here", to which a colleague responded, "Yes, but this is above our heads". There

is little evidence that policy addresses any issues of changing power dynamics. We cannot ascertain from the material whether the silence on the policy level is a conscious decision or these potential consequences have simply not been considered.

The Norwegian public sector operates on the basis that innovation happens at the sectoral and organisational level (Difi 2018). In the data-driven context, this translates to central initiatives around the facilitation of data sharing, but data usage remains the responsibility of the local level. Our findings show that, with the exception of the AI forum, activities are indeed happening locally, with no central coordination. There is currently no way of knowing what data-driven projects are being planned or, indeed, in production, which, as Brauneis and Goodman (2018) write, is a major transparency concern. That data-driven practices may have unintended societal consequences was recognised but not problematised by most practitioners. This is due largely to the fact that their individual projects may indeed be innocuous and have minimal societal impact; when combined, however, "small" and fragmented initiatives may actually have a real impact on the state-citizen relationship. The majority of current projects are based on control. Each of these can be justified from an organisational perspective, but, when taken on a national level, the question needs to be asked of whether this is moving in the direction of better services to citizens, as envisioned, or could signify a shift towards more state control.

A data-driven public administration brings to the fore many of the classic questions of public administration that are tied to equity, accountability, political legitimacy and what it means to be a professional public administrator, challenging them in fundamental ways (Barth and Arnold 1999; Bullock 2019; Veale, Van Kleek and Binns 2018). There is no evidence in the data material that they are being seriously considered by any stakeholders. For example, questions of human agency in decision-making (Lipsky 1980) were rarely mentioned, although most projects are expected to replace (parts) of discretional decision-making. Again, data-driven public administration is seen mainly as a technical or organisational issue hindered by existing legal frameworks. Many decisions around the implementation of data-driven public administration are delegated to data scientists, information management specialists and architects, who do not have the expertise to understand the potential implications for governance of the administrative state (Barth and Arnold 2019). Coupled with this, the neutral language of technology in which data-driven public administration is framed often facilitates designers' neglecting the processes of democracy and accountability (Veale, Van Kleek and Binns 2018).

The discursive context of data-driven public administration is significantly shaped by corporate technology companies (Andrews 2019). Although most public sector entities in Norway encourage the in-housing of competence, the growing reliance on private sector infrastructure, such as Microsoft Azure and Amazon web services, is an unproblematised issue in the discourse. Data-driven public administration might further intertwine the public and business spheres and give the private sector increasing control over public sector infrastructure and data (Redden 2018). This is further compounded by our observation that citizens are barely included in processes at either the policy or the practitioner level. The discourse is reserved for politicians, private and public sector

'experts' and officials, despite research finding the immense importance of popular support in administrative reform (Caiden 1999). There is a widespread impression that the Norwegian population has high levels of digital literacy and is therefore able to grasp, use and assess data-driven tools. Digital competence measures, however, do not measure data literacy, as only the use, knowledge and command of digital services is measured (Kompetanse Norge 2018). The general competence related to data-driven technology in the population is low, which makes it difficult to initiate public discussion, allowing greater room for influence by corporate interests.

Few practitioners regarded ethics as a major challenge to act upon. This result was somewhat surprising given the popularity of ethics in this domain. As with other aspects of this analysis, a nuanced picture emerged. When challenged, many had just started working and had not yet encountered any ethical dilemmas, which often surface further down the road of development and implementation (Veale, Van Kleek and Binns 2018). While public sector entities with sensitive data (such as health data) view ethics as a major challenge, those with more technical or non-sensitive data see this as little or no challenge. Again, practitioners seem to believe that, once privacy and legislation are respected, the solution is automatically ethical. The current framing of the problem around personal data and privacy leaves little room for consideration of the potential impact of nonpersonal, synthesised and anonymised data. This type of data operates outside the scope of data protection law (Andrew and Baker 2019) and is largely considered benign. However, there is a growing body of research pointing to the concern that this data can still have a major societal impact. Anonymised and aggregated data can still be sensitive and political (Kitchin 2014), and, even when anonymised, behavioural data can have immense power to influence and discriminate (Zuboff 2019). The fact that many governments participate in the sharing and utilisation of this type of data (Andrew and Baker 2019), to which Norway is no exception, suggests the need to incorporate this perspective into the discourse.

Most administrative reforms fail, as reformers are often too optimistic and unrealistic, falling into the many traps of implementation (Caiden 1999). Our findings at the practitioner level echo those of Hagendorff and Wezel (2019), who point out that, although data-driven technologies draw on a mythical character, they still require a significant amount of hands-on work and produce a variety of hidden costs. The challenges of implementation presented in this paper are interwoven and mutually dependent. Acting upon one challenge will not solve all the others; in fact, it might elicit other, unforeseen consequences. However, the public sector and policy discourse are often concerned with challenges in isolation rather than with interdependent issues to consider. These findings therefore challenge the often-deterministic check box approach as embraced by policymakers, industry and practitioners. Data-driven practices are nonlinear and ambiguous, administrative reform a dynamic process. As the installed base of public administration is fragmented and varies highly within the sector, this indicates that there is no one-size-fits-all solution to the challenges that public entities face. The sector's work itself is both enabled and limited by the bottom line of creating public value and public mandates. One might argue that these challenges are unseen because their implications are not immediately

visible, particularly when social scientists, public administration scholars and citizens are not involved.

Conclusion

This paper asks the overall research question of what challenges are encountered and problematised in a nascent phase of data-driven public administration implementation. Studying the practical experience of implementing administrative reform and grasping the challenges of contemporary government provide practitioners, researchers and policymakers with "real-world" experience of the grand idea of data-driven public administration and help to root discussions about the "what", "how" and "where to" within the setting of public administration. There is a distinct lack of research in this area. It is crucial to increase this in order to maintain integrity in what is a paradigm shift for public administration and to expand research beyond the UK and US contexts. By laying bare the complexity involved in data-driven public administration implementation, we endeavour to whet the appetite of public administration scholars to engage more deeply and to provide insights for policymakers and public servants alike, which heretofore may not have been visible.

Keeping track of the ongoing data-driven transformation of society, determining its potential social implications and finding appropriate social and legal responses prove to be challenging (Kitchin 2014). This paper adopts a practice approach to the phenomenon, focusing on both the institutions and practitioners currently working on its realisation in the Norwegian context. Highlighting the challenges requiring action as identified by practitioners, we contextualise the experience of public sector actors within policy and research in the field. We discuss unseen issues that are simply either not on the radar or considered inconsequential. Setting this ambiguity tied to data-driven practices within the broader policy and research context draws a complex picture of technical, organisational, regulatory and cultural issues, which bears much resemblance to earlier research on administrative reform. What we can observe is that, when embarking on their journey to deploy these technologies to access and utilise the "goldmine" of data, deterministic views and hype tied to data-driven practices at the policy level often fall apart when applied to political, noisy, stressful, complex and contested deployment settings, such as public administration (Veale, Van Kleek and Binns 2018).

Understanding the interplay of both seen and unseen challenges and the practical experiences of public sector practitioners can contribute to a broader understanding of the phenomenon. A holistic approach at a political, administrative and societal level will help to frame the discussion and broaden the perspective beyond the current focus. The empirical account given here does, though, also have theoretical implications for the field of public administration reform. The limitations of this study include its small sample size (due to few organisations having embarked on the data-driven journey), and it does not follow the cross-implementation process over time. Implementation is highly context dependent, as shown in this analysis; the findings therefore cannot be generalised for all public sector reforms; however, they provide an interesting starting point for further research. The unseen issues as discussed in this paper

and the contested concept of the data-driven public sector are particularly ripe for further research. It is beyond the scope of this paper to evaluate success or failure; what is clear, however, is that implementation approaches currently struggle to understand and appreciate the complexity of the challenge. Implementation is a long, arduous and uncertain process (DeLeon 1999); however, given that this is still in the nascent phase, there is time to adjust the course.

Acknowledgements

The authors would like to thank Alice Elizabeth Dalsjø for her research assistance transcribing interviews and sorting literature. In addition, we would like to thank Petter Almklov, Endre Grøtnes, Arild Jansen, Malcom Langford, Emil Røyrvik, Dag Wiese Schartum and Martin Standley for reading and commenting on earlier drafts of the article. We would also like to thank the Norwegian AI forum for their assistance distributing the survey.

Conflict of Interest

No conflict of interest has been declared by the authors.

Funding

The research presented in this paper was partly funded by the Norwegian University of Science and Technology´s (NTNU) digital transformation initiative project "Digital infrastructures and citizen empowerment" (DICE) and the Norwegian Research council's public administration PhD program.

References

Agarwal, P. K. (2018) Public administration challenges in the world of AI and bots, *Public Administration Review*, 78 (6): 917-921.

Alston, P. (2019) *Report of the Special Rapporteur on Extreme Poverty and Human Rights.* UN document. A/74/48037.

Andrew, J. & M. Baker (2019) The general data protection regulation in the age of surveillance capitalism, *Journal of Business Ethics*,

Andrews, L. (2019) Public administration, public leadership and the construction of public value in the age of the algorithm and 'big data', *Public Administration*, 97: 296-310.

Barth, T.-J. & E. Arnold (1999) Artificial intelligence and administrative discretion: Implications for public administration, *American Review of Public Administration*, 29 (4): 332-351.

Beer, D. (2017) The social power of algorithms, *Information, Communication & Society*, 20 (1): 1-13.

boyd, D., & K. Crawford (2012) Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon, *Information, Communication & Society*, 15 (5): 662-679.

Brauneis, R., & E. P. Goodman (2018) Algorithmic transparency for the smart city, *Yale Journal off Law and Technology*, 20: 103-176.

Bullock, J. B. (2019) Artificial intelligence, discretion, and bureaucracy, *American Review of Public Administration*, 49 (7): 751-761.

Busch, P. A. & H. Z. Henriksen (2018) Digital discretion: A systematic literature review of ICT and street-level discretion, *Information Polity*, 23 (1): 3-28.

Caiden, G. E. (1999) Administrative reform – Proceed with caution, *International Journal of Public Administration*, 22 (6): 815-832.

Caiden, G. E. & P. S. Puniha (2011) Putting public governance innovation into perspective: From administrative reform to innovation discourse, *Innovation and the Public Sector*, 15: 23-38.

Creswell, J. W. & V. L. P. Clark (2011) *Designing and Conducting Mixed Methods Research*, Sage Publications, London.

De Leon, P. (1999) The missing link revisited: Contemporary implementation research, *Review of Policy Research*, 16 (3-4): 311-338.

De Leon, P. & L. De Leon (2002) What ever happened to policy implementation? An alternative approach, *Journal of Public Administration Research and Theory*, 12 (4): 467-492.

Dencik, L. (2019) 'Situating practices in datafication—From above and below' in H. Stephansen & E. Treré (eds.), *Citizen Media and Practices*, Routledge, London and New York.

Desrosières, A. (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*, Harvard University Press, Cambridge.

Difi (2018) *Deling av data: Konseptvalgutredning*. Retrieved from https://www.difi.no/sites/difino/files/deling_av_data_kvu_sladdet.pdf.

Dille, T & J. Söderlund (2013). Managing temporal misfits in institutional environments: A study of critical incidents in a complex public prosject, International Journal of Managing Projects in Business,2 6(3):552-575

Dølvik, J. E., T. Fløtten, J. M. Hippe & B. Jordfald (2014) Den nordiske modellen mot 2030. Et nytt kapittel? (Fafo-rapport 2014: 46). Retrieved from https://www.fafo.no/index.php/zoo-publikasjoner/fafo-rapporter/item/den-nordiske-modellen-mot-2030-et-nytt-kapittel

Dunleavy, P., H. Margetts, S. Bastow & J. Tinkler (2005) New public management is dead—Long live digital-era governance, *Journal of Public Administration Research and Theory*, 16: 467-494.

Elish, M. C. & d boyd (2018) Situating methods in the magic of big data and AI, *Communication Monographs*, 85 (1): 57-80.

Fredriksson, C., F. Mubarak, M. Tuohimaa & M. Zhan (2017) Big data in the public sector: A systematic literature review, *Scandinavian Journal of Public Administration*, 21 (3): 39-62.

Hagendorff, T. & K. Wezel (2019) 15 challenges for AI: Or what AI (currently) can't do, *AI & Society*, 35: 355-365.

Henley, J. & R. Booth (2020, Feb. 5) Welfare surveillance system violates human rights, Dutch court rules. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules.

Hill, M. & P. Hupe (2014) *Implementing Public Policy: An Introduction to the Study of Operational Governance*, Sage, London.

Hintz, A, L. Dencik &K. Wahl-Jorgensen (2019) Digital Citizenship in a Datafied Society. Cambridge: Polity Press

Hovde Lyngstad, T. & T. Skardhamar (2011) Nordic register data and their untapped potential for criminological knowledge, *Crime and Justice*, 40 (1): 613-645.

Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*, Sage Publications, London.

Klievink, B., B. J. Romijn, S. Cunningham & H. de Bruijn (2017) Big data in the public sector: Uncertainties and readiness, *Information Systems Frontiers*, 19 (2), 267-283.

Kommunal- og moderniseringsdepartementet (2019) En digital offentlig sektor: Digitaliseringsstrategi for offentlig sektor 2019–2025. Retrieved from https://www.regjeringen.no/no/dokumenter/en-digital-offentlig-sektor/id2653874/.

Kompetanse Norge (2018) Grunnleggende digitale ferdigheter. Retrieved from https://www.kompetansenorge.no/statistikk-og-analyse/grunnleggende-digital-ferdigheter/

Kramer, K. L. & L. King (2006) Information technology and administrative reform: Will the time after eGovernment be different?, *International Journal of Electronic Government Research*, 2 (1): 1-20.

Lillington, K. (2020) Leaving cert: Why the government deserves an F for algorithms. Retrieved from https://www.irishtimes.com/business/technology/leaving-cert-why-the-government-deserves-an-f-for-algorithms-1.4374801.

Lipsky, M (1980) Street-Level Bureaucracy: Dilemmas of the Individual in Public Services, Russell Sage Foundation, New York.

Luthfi, A. & M. Janssen (2019) Open data for evidence-based decision-making: Data-driven government resulting in uncertainty and polarization, *International Journal on Advanced Science, Engineering and Information Technology*, 9 (3): 1071-1078.

Maciejewski, M. (2016) To do more, better, faster and more cheaply: Using big data in public administration, *International Review of Administrative Sciences,* 83 (1): 120-135.

Margetts, H. (2009) Public management change and e-government: The emergence of digital-era governancein Chadwick A. & P. N. Howard, *Routledge Handbook of Internet Politics*, Routledge, London and New York

Mergel, I., R. K. Rethemeyer & K. Isett (2016) Big data in public affairs, *Public Administration Review,* 76 (6): 928-937.

Misuraca, G. & C. van Noordt (2020) Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg.

Muid, C. (1994) Information systems and new public management—A view from the centre, *Public Administration*, 72: 113-125.

NOU 1988:40 (1988) Datapolitikk i staten i 1990-årene. Retrieved from https://www.regjeringen.no/no/dokument/nou-ar/nou-samandrag.

OECD (2017) *Digital Government Review of Norway—Boosting the Digital Transformation of the Public Sector*. Retrieved from https://www.oecd.org/gov/digital-government/digital-government-review-norway-recommendations.pdf.

O'Neil, C. (2016) Weapons of math destruction. How bog data increases inequality and threatens democracy. d: Portlanoradway Books

O'Toole, L. J. (2000) Research on policy implementation: An assessment and prospects, *Journal of Public Administration Research and Theory*, 10 (2): 263-288.

Pasquale, F. (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge.

Pencheva, I., M. Esteve & S. J. Mikhaylov (2018) Big Data and AI – A transformational shift for government: So, what next for research?, *Public Policy and Administration*, 35 (1): 24-44.

Redden, J. (2018) Democratic governance in an age of datafication: Lessons from mapping government discourses and practices, *Big Data & Society*, 5 (2): 1-13.

Ruppert, E., E. Isin & D. Bigo (2017) Data politics, *Big Data & Society*, 4 (2):1-7.

Schei, A. (2020, July 15) Kunnskapsdepartementet ber IB rydde opp i karakterrot. Retrieved from https://khrono.no/kunnskapsdepartementet-ber-ib-rydde-opp-i-karakterrot/502707.

Sun, T. Q. & R. Medaglia (2019) Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare, *Government Information Quarterly*, 36 (2), 368-383.

Tupasela, A., K. Snell & H. Tarkkala (2020) The Nordic data imaginary, *Big Data & Society*, 7 (1): 1-13.

Ubaldi, B., F. Gonzalez-Zapata & M. P. Barbieri (2020) *Digital Government Index 2019 Results*. Retrieved from http://www.oecd.org/gov/digital-government-index-4de9f5bb-en.htm.

van Dijck, J. (2014) Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology, *Surveillance & Society*, 12 (2): 197-208.

van Ooijen, C., B. Ubaldi & B. Welby (2019) A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance, *OECD Working Papers on Public Governance,* 33, OECD Publishing.

Veale, M., M. Van Kleek & R. Binns (2018) Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*: 440.

Wagenaar, H. & M. Wood (2018) The precarious politics of public innovation, *Politics and Governance*, 6 (1): 150-160.

Williamson, B. (2014) Knowing public services: Cross-sector intermediaries and algorithmic governance in public sector reform, *Public Policy and Administration*, 29 (4): 292-312.

Wirtz, B. W., J. C. Weyerer & C. Geyer (2019) Artificial intelligence and the public sector—Applications and challenges, *International Journal of Public Administration*, 42 (7), 596-615.

Yeung, K. (2018) Algorithmic regulation: A critical interrogation, *Regulation & Governance*, 12: 505-523.

Zuboff, S. (2019) *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*, Profile Books, London.

Appendix A: Survey template

1. In which organisation are you employed? [Free text question]

2. What job code/title do you have? [Free text question]

3. In your opinion, how data-driven is your organisation in relation to other actors in the public sector? Please answer this on a scale of 1 to 10, with 10 being very data driven and 1 being minimally data driven. [Likert scale]

4. In your opinion, how highly prioritised is AI/data science in your organisation? Please answer this on a scale of 1 to 10, with 10 being highly prioritised and 1 being not prioritised. [Likert scale]

5. How is the work with AI/data science organised within your organisation? Here, examples of answers are (not exhaustive): the IT department; the statistics/analysis department; integrated across the organisation. [Free text question]

6. Who are the main AI model developers for your organisation? [Drop-down list]

- *a. Mainly in-house staff*
- *b. Approximately 50/50 consultants and staff*
- *c. Mainly hired consultants*
- *d. Other*

7. How far has your organisation progressed in the work with AI/data science? [Drop-down list]

- *a. Starting to think about using it*
- *b. Planning phase*
- *c. Start-up phase*
- *d. Testing*
- *e. Production*
- *f. Operation/management*

8. Follow up to 7: (a) What are you considering using AI/data science for? [Free text question]; (b) What are you planning to use AI/data science for? [Free text question]; (c) What is AI/data science used for in your organisation? [Free text question]

9. Below are a number of potential reasons for using AI/data science. Please rank them from most important (highest) to least important (bottom). [Ranking]

- *a. More efficient decision-making processes*
- *b. Better quality and timeliness in decisions*
- *c. More precise predictions*

d. Increased user orientation
e. Increased innovation and business development
f. Increased employee satisfaction
g. Reduced costs

10. Below are a number of potential challenges around the adoption of AI in the public sector. Please rate how large the challenge is for your organisation on a scale from 1 to 5, with 1 being no challenge and 5 being a very large challenge. [Likert scale]

a. Technical infrastructure
b. Ethics
c. Organisational culture
d. Access to data
e. Data quality
f. Privacy and security
g. Analytics competence
h. Legal and regulatory framework
i. Citizens' insecurity and willingness to accept AI/data use
j. Uncertainty around what AI could be used for in the organisation
k. Lack of funding
l. Pressure from management to deliver
m. Fear of downsizing in the organisation

11. Are there other challenges when using AI/data science in your organisation? [Free text question]

12. Do you have a project/activity in the AI/data science area that you would like to tell us about so that we can share it with others? Please give a short description of the project/activity. [Free text question]

13. Would you or anyone else in your organisation be willing to be interviewed by us? We are looking for insights into what the public sector needs in the AI/data science field. [Yes/No]

a. Contact information [Free text question]

14. Do you have anything else you would like to add or comments on this questionnaire? [Free text question]

Appendix B: Interview guide

The guide followed a general approach, with a number of common questions posed to each interviewee based on the survey questions, which are found below. An individual guide was, however, prepared for each organisation, adapted to the responses that were made in the survey. Here, we made comments on interesting issues to follow up with the interviewee(s). We presented the responses to the interviewees and encouraged them to elaborate and explain their justifications for the responses.

Questions

1. What is your name and job title, and what are you working with in your organisation?

2. In your opinion, how data-driven is your organisation in relation to other actors in the public sector?

a. Why did you position your organisation here?
b. What does it mean to your organisation to become data driven?

3. In your opinion, how highly prioritised is AI/data science in your organisation?

4. How is the work with AI/data science organised within your organisation?

a. Why is the work organised this way?

5. Who are the main AI model developers for your organisation?

a. Why did the organisation choose to place the responsibility here?

6. How far has your organisation progressed in the work with AI/data science?

a. What are you currently planning to use this technology for, or what is it already used for?
b. Do you have a project/activity in the AI/data science area that you would like to tell us about?

7. In the survey, you ranked potential reasons for using AI/data science in your organisation. Can you elaborate on this ranking?

a. Are there any other reasons that you would like to add?

8. Graph: Here we presented the interviewee(s) with the graph of their results measured against the average of the AI forum and encouraged them to discuss each of the challenges and justifications as to their ranking. They were asked to elaborate on each of the challenges.

a. Are there other challenges when using AI/data science in your organisation?
b. Would you like to discuss challenges in a specific project?

9. Is there anything else you want to discuss or bring up?

Notes

1. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2634733

[Paper 3]
In Search of the Citizen in Public Administration Datafication

This paper is co-authored with Heather Broomfield and published in *Big Data & Society* (2022) 9(1):1-14.

Original Research Article

BIG DATA & SOCIETY

Big Data & Society
January–June: 1–14

DOI: 10.1177/20539517221089302
journals.sagepub.com/home/bds
SAGE

In search of the citizen in the datafication of public administration

Heather Broomfield[1] and **Lisa Reutter**[2]

Abstract

The administrative reform of the datafied public administration places great emphasis on the classification, control, and prediction of citizen behavior and therefore has the potential to significantly impact citizen–state relations. There is a growing body of literature on data-oriented activism which aims to resist and counteract existing harmful data practices. However, little is known about the processes, policies, and political-economic structures that make datafication possible. There is a distinct research gap on situated and context-specific empirical research, which critically interrogates the premises, interests, and agendas of data-driven public administration and how stakeholders can impact them. This paper therefore studies the conditions of participation in public administration datafication. It asks the overall research question of how citizens are problematized and included in policy and practitioner discourse in the datafication of public administration. The paper takes Norway as its case and applies Cardullo and Kitchin's scaffold of smart citizen participation at the system level. It makes use of a unique empirical insight into the field, consisting of a survey, interviews, and an extensive document analysis. Unexpectedly, we find that citizens and civil society are rarely engaged in this administrative reform. Instead, we identify a paternalistic, top-down, technocratic approach where the context, values, and agendas of datafication are obscured from the citizen.

Keywords

Public administration, datafication, citizen participation, civil society, Norway, artificial intelligence

Introduction

Citizens are increasingly faced with datafied services in their interactions with public administration (Misuraca and van Noordt, 2020). This paradigm shift in the public sector is leading to profound changes in the way modern public administration learns about, engages with, and responds to citizens (Redden, 2018; Dencik et al., 2019). This paper asks the overall research question: *How are citizen perspectives problematized and included in policy and practitioner discourse in the datafication of public administration*? It examines both *who was consulted* in the policy-making process guiding this administrative reform and *the discourse within the resulting policy and among practitioners in the field*. Taking Norway as our case, we perform an empirical investigation to obtain a better understanding of how, where, and when citizens are included in this process. We apply a multi-method approach, which includes a content analysis of key policy documents and a secondary analysis of data from surveys and interviews with practitioners. Datafication research, as with administrative reform itself, is still in a nascent phase in both Norway and beyond (Broomfield and Reutter, 2021). This provides an opportunity to investigate datafication in the making and contribute to the critical understanding of who datafication serves, what is intended to be optimized and for whom, and who gets to decide (Ruppert et al., 2019; Zuboff, 2019; Crawford, 2021).

The datafied and disempowered citizen is an important object of investigation (Gabrys, 2019; Hintz et al., 2019). Early research on the role of citizens in digitalization has been predominantly framed around the idea of citizen empowerment, due to increased possibilities for participation and the enabling of new interactions between citizens, and citizens and the state (Mossberger et al., 2007; Chun et al., 2010). However, this research fails to grasp the complexity of how data are used to categorize, classify, and

[1]Public and International Law, University of Oslo, Oslo, Norway
[2]Department of Sociology and Political Science, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding author:
Lisa Reutter, Norwegian University of Science and Technology, Trondheim, Norway.
Email: lisa.m.reutter@ntnu.no

profile citizens' activities and behavior (Hintz et al., 2019). Data activism studies have come to the fore in recent years. This field investigates how citizens react to and resist problematic data practices (Meng and DiSalvo, 2018; Lehtiniemi and Ruckenstein, 2019), rather than studying the conditions of participation in the production and implementation of datafication. To understand the emergence of the datafied citizen, Barassi (2019: 426) stresses that "we have much to gain if we focus on process, on the multiple ways in which individuals are being turned into datafied citizens, on the policies and political-economic structures that make this datafication possible." Consideration is needed around how state–citizen relations are configured through datafication with the premises, interests, and agendas demanding to be critically interrogated in context (Dencik et al., 2019). Critical data studies have paid scant attention to the experiences and negotiations of non-expert citizens living with data (Kennedy et al., 2020). Some empirical research is emerging, such as that by van Zoonen (2020), who found in a case study of Dutch municipalities that datafication often escapes democratic decision-making, and that citizens as key stakeholders are not actively informed, let alone invited to participate. This paper contributes to this literature.

Norway is a particularly interesting case to investigate. Datafication lies at the heart of the future imagery of the Nordic welfare state (Kaun and Dencik, 2020). The Norwegian government has collected vast amounts of data on the population for decades, and it is stressing the enormous value and untapped potential of this perceived goldmine (Difi-rapport, 2018: 7). "Data is the new oil" has become a powerful metaphor in a country currently in search of something to replace its massive oil sector. There is widespread trust in government, which "intersects with the popular belief that technological progress is inevitable, apolitical, and problem free" (Sandvik, 2020: 2). Participation is also a fundamental component of the Norwegian welfare state. Corporatist and consensual decision-making, which relies upon the inclusion of societal stakeholders, in addition to academics, public servants, and the private sector, is deemed to be a key element in the construction and implementation of policy (Christiansen et al., 2010).

A recent study of the participatory turn in Norwegian public administration has found clear indications of citizen inclusion and grassroots involvement in climate change, energy, family, and gender policy fields (Krick et al., 2019). One would expect this to be replicated in the datafication field. However, during a preliminary analysis of our data, we found the citizen to be virtually absent. This triggered a secondary analysis to study how citizen perspectives are problematized and included in policy and practitioner discourse—from a system-level perspective. We situate the case within Cardullo and Kitchin's (2019) scaffold of smart citizen participation, as it provides an analytical framework to discuss citizens' roles, the nature of their involvement, and the underlying discourse at several levels of public administration. Our findings expose an unexpected dearth of citizen participation in both the planning and implementation of the datafied public administration in Norway.

Prior to introducing the analytical framework and research design of the paper, we further elaborate on the concepts of a datafied public administration and citizen participation. This allows us to frame the problem and highlight what is at stake. The analysis then provides empirical insights into the inner workings of public sector datafication. This paper then culminates in a discussion of the role and agency of citizens in this process.

Background and previous research: data-driven public administration and citizen participation

The production and recirculation of data interests those who exercise power. It is therefore irresistible to the modern state and its executive powers of public administration (Ruppert et al., 2019). The modern state and data are already inseparably interwoven, as the availability of statistical information is a condition and necessity for any democracy (Desrosières, 1998). Data accumulation and recirculation in ever more aspects of citizens' lives beyond statistics is, however, a more recent phenomenon. The logic of datafication is envisioned to be integrated as a key component of how decisions about citizens are made across their social, political, economic, and cultural participation, disrupting the social world in all its forms (Kennedy, 2018; Hintz et al., 2019). While digitalization refers to a process of converting analog information into binary code so it can be processed by computers, datafication describes the process of quantifying every aspect of the world so that it can be analyzed (Mayer-Schönberger and Cukier, 2013). Public administration datafication aims to promote the idea of data as an asset to be integrated into policy making, service delivery, organizational management, and innovation (van Ooijen et al., 2019).

Datafication is predominantly associated with big-tech; however, the powerful and pervasive imaginaries associated with it penetrate both the private and public sector (MayerMayer-Schönberger and Cukier, 2013). This paper situates datafication beyond big-tech and within the realm of public administration. It regards datafication as a sociotechnical process characterized by the ever-growing utilization of advanced methods to analyze and recirculate data (Kitchin, 2017). Datafication operates with the logic of categorization, classification, scoring, and selecting (Dencik et al., 2019). Recirculating government data is a prerequisite. Marrying public and private data is a further ambition. This is both an extension of previous data practices and a

profound change, as data is now imagined to drive decision-making in all aspects of public administration in ever more complex ways. It is built on the overall ideological foundation of dataism—a belief in data as the enabler of a better, more effective, and objective society (van Dijk, 2014). This is a paradigm shift, bringing with it a new form of knowledge production (Yeung, 2020). While the concept of New Public Management (NPM) focused on the integration of private sector management ideas into public administration, datafication is primarily concerned with a reform of knowledge production practices and algorithmic forms of ordering. Huge amounts of data and complex analysis promise a supposed ability to reveal a hidden mathematical order to the world that is superior to direct experiences (McQuillan, 2018; Dencik et al., 2019).

The datafied public sector is far from a benign instrument, leading to the grand realization of the ever-elusive efficient and effective public administration. Citizens are increasingly surveilled and governed, with data speaking for and about individual lives, extending beyond individual choice or control (Barassi, 2019). Data collection, storage, retrieval, and analysis make objects and subjects visible, enhancing the ability of public administration to understand, predict, and control (Hintz et al., 2019). Providing public administration with a golden view of society and empowering the state (Dencik et al., 2019). Data are not mere representations but instead generative of new power relations, which "may seriously erode capacity for democratic participation and individual flourishing" (Yeung, 2017: 119).

Citizen participation in public administration is a well-established research field that studies how democratic societies have increasingly involved citizens in policy production and implementation (Roberts, 2004; Bingham et al., 2005). Some researchers purport that multi-actor collaboration in the form of co-creation is filling the void left after the demise of NPM in Europe (Ramaswamy and Ozcan, 2014; Torfing et al., 2019). As Mellouli et al. (2014) pointed out, smart government requires engagement with citizens and interactive processes, both among citizens and between citizens and government, to create and implement public policies and decisions in a transparent and responsible manner. Concern is, however, voiced about data imaginaries and the role of citizens in Nordic data visions (Tupasela et al., 2020). Healthy state–citizen relations in democracies rely on the ability of citizens to scrutinize and challenge public sector decisions and processes, which according to Kaun and Dencik (2020) is increasingly threatened by datafication. There are several studies on participation in smart city projects (e.g. Benouaret et al., 2013; Berntzen and Johannessen, 2016); however, a distinct research gap remains in the broader context of datafication of public administration.

Citizens are increasingly concerned with how data are used and in which contexts (Kennedy et al., 2020). There is a growing body of literature around data activism where citizens react to and resist harmful data practices (Milan, 2016). This form of activism is primarily reactive, based on bottom-up initiatives largely concerned with how citizens can influence existing systems that have already produced negative outcomes (Beraldo and Milan, 2019). It also relies on technology and data experts and therefore tends to be dominated by elites (Kennedy, 2018). Critical data studies have paid scant attention to the experiences and negotiations of non-expert citizens living with data (Kennedy et al., 2020). One exception is Van Zoonen (2020), who, in her study on Dutch municipalities, found that "despite the municipal goals of dialogue and self-direction, a top-down practice of control is emerging, which has its roots in the notion of 'data steering' and which leaves no space for citizens or other stakeholders" (2020: 4). As datafication has the potential to impact state–citizen relations, it is vital to investigate how both non-expert citizens and their representatives in the form of civil society, as key stakeholders, are envisioned by and engaged with during the entire process—from policy production to implementation.

Analytical framework

This paper studies the degree to which citizens are included and able to challenge the premises, interests, and agendas of datafication (Dencik et al., 2019) by (a) examining *who was consulted* in the policy-making process and (b) *the discourse within resulting policy and among practitioners in the field*. A number of models across different disciplines could be used to assess and investigate these questions, including analysis of spaces of participation (Cornwall, 2002) and investigations of participation as communicative action, where citizens' understanding of their decision-making power is assessed (Chang and Jacobson, 2010). As this paper requires an analytical framework that encompasses both the nature of citizen involvement itself and problematizations, we have chosen to situate the Norwegian public sector datafication process within Cardullo and Kitchin's (2019) scaffold of smart citizen participation. This builds upon Arnstein's (1969) ladder of participation, providing a typology of citizen roles, the form and nature of citizen involvement, and the political discourse underlying datafication efforts. We treat participation as a situated practice, framing citizens' possibilities with reference to actual political, social, and practical particularities rather than idealized notions of democratic practice (Cornwall, 2002).

Cardullo and Kitchin's (2019) scaffold provides this paper with a candid analytical base from which to investigate and discuss the conditions of participation in the datafication of the Norwegian public administration. We adapt the scaffold by applying it to study different levels of public administration. Rather than concentrating on specific projects, we investigate how citizens are included and problematized

Table 1. Adapted and condensed scaffold of smart citizen participation (Cardullo and Kitchin, 2019).

Form and level of participation		Political discourse/ framing	Modality
Citizen power	Citizen control	Rights, social/ political citizenship, commons	Bottom-up
	Delegated power		
	Partnership	Participation, co-creation	
Tokenism	Placation		Top-down
	Consultation	Civic engagement	
	Information		
Consumerism	Choice	Capitalism, market	
Non-participation	Therapy	Stewardship, technocracy, paternalism	
	Manipulation		

during the policy production process and the implementation phase. This allows us to produce a contextualized critique of an administrative reform that is circulated and enacted throughout the public administration, as recommended by both Barassi (2019) and Dencik et al. (2019).

The scaffold (Table 1) has four levels of participation, as elaborated by Cardullo and Kitchin in the following ways.

First, **non-participation** affords citizens limited rights and possibilities for changing outcomes. Instead, they are nudged and steered toward certain behaviors, perceived as users, patients, and learners, and reduced to data points. Services are delivered on their behalf and supported by a strong technocratic paternalistic impulse. Citizens are neither included in nor consulted about how datafication processes are formulated, produced, and deployed. They are often subjected to new forms of governance that further dissolve transparent and democratic processes, due largely to the proliferation of hybrid configurations in the form of public–private agencies, new administrative units, and "experts," which often operate beyond the state.

Second, **consumerism** encompasses the idea of the citizen as a consumer with a restricted range of control, selecting from a marketplace of a limited number of predetermined services. If citizens are consulted, it is in the form of feedback to specific pilots or products, thus tweaking existing designs. Consumerism is enabled by strong technocratic framing.

Third, **tokenism** consists of a lower form of sharing, informing, and shaping decision-making, intended to create transparency and accountability. Information is, however, predominantly unidirectional and shared only after key decisions are made. In its higher form, tokenism consists of placation and consultation. Citizens are given a voice via active feedback and participation in user testing. Placation is then about citizens suggesting alternatives with some ability to reshape plans and actions. Tokenism largely reproduces the dominant interests of public administration and follows a predetermined course.

Fourth, **citizen power** is about the redistribution of power from those in control to citizens, where communities can negotiate and engage in datafication processes during the planning and execution phase, and citizens are provided with platforms to resolve differences.

These concepts form the basis of our analysis.

Methodology

A situated analysis of datafication allows researchers to understand underlying social mechanisms as well as the imaginaries of agents responsible for datafication (Dencik, 2020). Using interviews, a survey, and document analysis, we adopted a multi-method approach to answering the overall research question of how citizens are included and problematized in policy creation processes and practitioner discourse in Norway (see Table 2).

The first part of the research consisted of an analysis of policy-making processes, which aimed to ascertain how national datafication policy has been produced and who was consulted in the process, thus investigating how citizens were included. The method section of each document and the accompanying official information on the relevant web pages were collected and analyzed. These documents were selected because they are either central national datafication policy documents or vital inputs to policy, providing the foundation from which datafication is framed and actualized in Norway. A list and description of the documents can be found in Appendix A.

The second part of the research investigated the discourse in the field to understand how citizens are problematized and envisioned. Both policy and practitioner discourses were central. Some of the material analyzed was initially collected to map the public sector's work with data-driven technology.[1] The empirical account given in this paper is based on an in-depth secondary analysis, which consists of a survey ($n = 35$) answered by practitioners in 26 public organizations and follow-up interviews with 12 entities (see Appendixes B and C). Practitioners who provided input to this study are system-level rather than street-level bureaucrats, as most activity can be observed here. The survey and interview sample consists of entities that operate at different administrative levels and vary in size, therefore presenting a diverse set of cases. Intelligence agencies were not included. Survey and interview responses were recruited from an informal public sector artificial intelligence (AI) practitioners' network. The aim of the initial study was to obtain an understanding of how different public sector entities engage in new data practices and the challenges that practitioners encounter. Respondees were not asked directly about how they included citizens in their work, as the data was not initially collected for the purpose of this paper. While it could

Table 2. Multi-method research design to study policy production process, and policy and practitioners' discourse.

Part 1: Process	Part 2: Discourse
Method section of strategy and policy documents ($n = 5$)	Survey ($n = 35$)
Official webpages describing policy production	Follow-up interviews with public entities ($n = 12$)
	Key policy documents and reports ($n = 5$)

be argued that this affects the reliability of the material, we believe that this strips the interviews of socio-political desirability, as the survey and interviews were intended to map the actual practice of datafication.

Part 2 of the research included an analysis of key policy documents. This paper made use of an explorative and abductive analysis process, going back and forth between data and literature, and adding new data to further extend our understanding of citizen participation. The analysis was conducted in three stages. Stage 1 consisted of a general content analysis of the interview material. All mentions of users, citizens, clients, or residents were identified and analyzed within the context. In addition, all survey answers were screened for these concepts. Upon discovering the scarcity of the terms citizen or residence, we extended the content analysis to policy documents in Stage 2. Stage 3 then combined all data and analyzed its content in relation to Cardullo and Kitchin's scaffold and the key concepts of user-centric and needs-based approaches.

The Norwegian context

The Norwegian welfare state is built on the principles of solidarity, equality, participatory democracy, and the protection of vulnerable citizens. Corporative pluralism, where collaboration with externals and interdependent decision-making with interest organizations and business representative organizations, is deemed fundamental to policy making (Rokkan, 1966). This model encourages the inclusion of civil society and organized interests in both policy-making and implementation. There are regularized procedures for public participation in the formulation of new measures. The scale, scope, and rigor of consultation between interest groups and the political executive is a distinguishing feature of decision-making in the Nordics (Arter, 2004). Hybrid advisory public committees' assembling a range of different agents, such as academics, stakeholders, and civil servants, are an important feature for conflict resolution, knowledge production, and "input democracy" (Christiansen et al., 2010). Corporatism has come under threat in recent decades due to a reduction in the number of corporative institutions and increased lobbyism in the political sphere, where private-sector interests dominate (NOU, 2003:19). A recent white paper on civil society (Meld. St. 10 (2018–2019) reaffirms and strengthens its role as a vital actor in public sector decision-making. It describes civil society as a direct participant in policy development, with a key role in protecting and strengthening civil rights and as a contributor to the development of norms and values, particularly when society is undergoing change.

The Nordics are regarded as data goldmines with massive amounts of high-quality granular data on citizens. This data collection practice emerged at the formation of the welfare state for decision-making purposes and to improve health and living conditions (Tupasela et al., 2020). There are hundreds of data registers, collecting data in areas such as tax, health, education, births, crime, and social security. Citizens are legally obliged to provide data. The personal number assigned either at birth or at the point of immigration follows individuals from the cradle to the grave (Frank, 2000). Most agencies use this number as the primary key to link datasets (Hovde Lyngstad and Skardhamar, 2011). Large amounts of "data exhaust" gleamed from digital interactions with the state are also collected. Public trust has given scope to expanding operations for data collection in recent decades (Tupasela et al., 2020). Data sharing is currently controlled by a strict and complex regulatory regime, with ongoing efforts to simplify this to break down the data silos and enable datafication.

The Norwegian public sector is large and fragmented. There are strict organizational, sectoral, and geographic boundaries where entities operate with a high degree of autonomy. A strategic body (SKATE) was founded to further public sector digitalization cooperation, and the Norwegian Digitalization Agency (Digdir) was established as a catalyst for the digitalization of the public sector. The responsibility for national digitalization policy lies with the Ministry of Local Government and Modernization (KMD), supported by Digdir.

Analysis: toward a datafied public administration

During a preliminary analysis of survey responses, we were intrigued by the discovery that practitioners ranked citizens' insecurity and willingness to accept AI/data use below all but one of the other challenges (see Figure 1). This provoked the question of how and when citizens are included in public administration datafication. Upon further investigation, we identified merely four mentions of the term "resident" or "citizen" in the interview material, contrasting with 79 occurrences of the terms "client/customer" or "user." This triggered our secondary analysis of the material and an investigation of policy and its production to ascertain the conditions of citizen participation. The results are presented below.

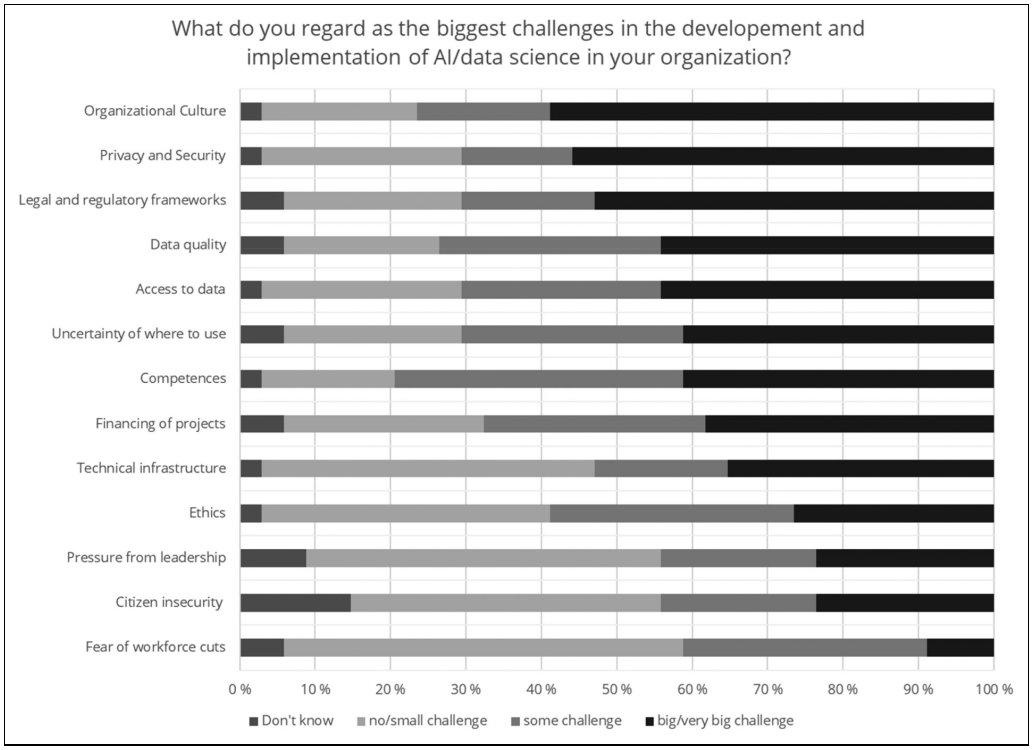


Figure 1. Perceived challenges to act upon, survey.

This is structured into three sections. First, we investigate the policy production process, asking who was consulted and exploring the main outcomes. Second, we focus on the discourse in policy and the practicing field, concentrating on the omnipresent concepts of user-centric and needs-based approaches. Finally, we analyze the idea of "doing the best for society," often referenced by practitioners. This analysis provides us with a basis from which to answer the research question of how citizen perspectives are problematized and included in policy and practitioner discourse.

Policy production and its output

Policy documents play a crucial role in laying the foundations for public sector datafication. The production process presents an intriguing object of investigation. Norwegian policy-making is deemed to be characterized by decision-making that is corporative, consensual, and integrative (Arter, 2004). In this section, we trace whether this is the case for the development of datafication policy.

The Digital Agenda of 2016 marked a paradigm shift when data recirculation became central to digitalization policy. An OECD report (2017) came hot on its heels, claiming the existence of untapped potential in Norwegian government data. These documents were instrumental in sparking the intensification of datafication in Norway, as evidenced by four major national datafication and digitalization policy actions in as many years. We investigate the process of producing these documents to determine how they were developed and who was engaged and consulted. Table 3 provides an overview of our findings.

The data sharing concept phase document (KVU) laid the foundation for many data policy prioritizations, particularly the digitalization strategy. It boasts of an extensive needs analysis to ensure that its proposed actions are rooted in user needs. Consultations were, however, limited to input from the private and public sector. We cannot find any attempt to engage directly with civil society or citizens. The process around the digitalization strategy paints a similar picture; four workshops for needs gathering were conducted, and over 300 people attended. There is no list of attendees publicly available; however, Difi states that "municipalities, county municipalities, state bodies and private companies from all over the

Table 3. Stakeholder engagement in Norwegian datafication policy-making.

2018	2019	2020	2021
Data sharing concept phase document (KVU)	**Public sector digitalization strategy 2019–2025**	**Artificial intelligence (AI) strategy**	**Whitepaper for data-driven economy and innovation**
Engagement with 50 stakeholders from the private and public sector.	4 workshops with 300 + attendees from the private and public sector.	11 events for the private sector, public sector, and technical AI research community; and 1 event organized by a trade union.	3 events targeted at the public sector, private sector, and education providers.
No open consultation	No open consultation	Open consultation, 51 submissions, 3 from citizen representatives, and 4 from trade unions.	Open consultation, 48 submissions, 2 from trade unions, 3 from civil society, and 2 from individuals.

country signed up" (Difi-rapport, 2018, 7). We found no reference to civil society involvement or any evidence that effort was made to engage with them as stakeholders.

In the AI strategy, we again found proclamations of an inclusive needs gathering phase. Events organized by the KMD were targeted at the public sector, the private sector, and the technical AI research community. One event was organized by a trade union. Workshops were supplemented by a formal open consultation. Analysis of the responses shows that citizen and employee representatives made up a mere 12% of the submissions. The consultation was in Norwegian, which would have hindered input from international civil society organizations active in this field.

The latest policy initiative is the White Paper for Data-Driven Economy and Innovation. It is premised on the assertion that data is a vital resource and must be further invested in as a source of innovation, new business models, and the improvement of existing digital services (KMD, 2021). The public sector, private sector, and education providers were invited to targeted events. The public consultation garnered a response pattern similar to the AI strategy. Two studies were commissioned: one from an expert group on "data sharing in the private sector," and the other from a law firm on whether there could be a legal requirement for the public sector to share data. The expert group was comprised solely of representatives from the private sector and technical AI research communities (Ekspertgruppen, 2020). The law firm, in agreement with the KMD, consulted exclusively with the private and public sectors (Kluge Advokatfirma, 2020: 7). Despite the flurry of activity in recent years, only one hybrid public advisory committee has been planned to provide input, despite these being central to Norwegian consensual democratic deliberation (Krick and Holst, 2021). This committee will look at data-sharing regulations, with the aim of increasing data sharing from the public sector (KMD, 2021: 69).

By analyzing discourse in documents, we found that datafication is framed in overwhelmingly positive terms, envisioned to improve public services and spur value creation in the private sector. It is deemed key to dealing with the growing complexity of society and a variety of economic and societal threats to the Norwegian welfare model. Economic assertions that the value of data is at least akin to that of oil, rely heavily upon a study (Menon, 2019), commissioned by the Confederation of Norwegian Enterprise (NHO), which is Norway's largest business representative body. As the following statements depict, there is a discernible sense of urgency, with data not just seen as a major driver for domestic economic growth and an antidote to shrinking public coffers, but also as the solution to ensuring sustainability, understanding demographics, and growing the green economy.

> We know that Norway will be affected by an aging population, climate change and increased globalization, and that we must work smarter and more efficiently to maintain competitiveness and the level of welfare in the years to come. Digitalization and new technology are the key to achieve this—and artificial intelligence will be central (KMD, 2020).

> The government wants Norway to take advantage of the opportunities inherent in data for increased value creation, more workplaces, and an efficient public sector. Better utilization of data is important if Norway is to succeed in the transition to a more sustainable society and a greener economy (KMD, 2021).

According to policy documents, datafication is both necessary and inevitable. Recirculation of vast amounts of data within and between the public and private sectors is seen as a prerequisite. Strengthening cooperation with the private sector to realize improved and more effective services and to lay the foundations for innovation is highly prioritized. The private sector is considered a partner to fulfill the roles of supplier of data-driven tools, co-creator of public services, user of public data, and data provider to the public sector. The private sector has been considerably involved in the policy development process.

Conversely, civil society and citizens are underrepresented in the material. The citizen fills one role, as a "user" of services, as discussed below.

User-centric and needs-based approaches

As the previous section shows, we failed to find evidence that insights into citizen needs and perspectives are produced through widespread engagement at the policy level. Yet citizens are presented as a demanding entity, apparently requesting faster, better, and more efficient services to be enabled by datafication. In this section, we concentrate on the discourse in policy and the practicing field, focusing on the omnipresent concepts of "user-centric and needs-based approaches." The term "user" is not reserved for citizens, as Norwegian policy defines users as "residents, voluntary, private, and public sectors" (KMD, 2019: 12).

User needs and placing the user at the center is seen as a major transformation for the Norwegian government and is a thread that runs through all digitalization policies since the Digital Agenda. The digitalization strategy states, "The user in the center is one of five key priorities in the Digital Agenda for Norway. The goal being that the user shall experience their encounter with the public sector to be coherent and effective, as one digital public sector" (KMD, 2019: 12). In this context, we can trace the origins of a discursive turn from citizen to user/customer, now pervasive throughout the sector and clearly apparent in the data. One interviewee stated that the focus on customers changes the way they work and refocuses attention toward the user experience. "If we say customers, we treat them a little differently […] we have a slightly different mindset. So, we say what can we do now? What kind of machine learning case can we implement to improve the customer experience for our customers?" (Informant C).

A valid question from which to start is: What needs are being prioritized and attributed to the "user" to justify datafication, and from where do they originate? Justification is based on the premise that users are frustrated by a fragmented digital public service offering, which is not on par with services from the private sector. The origin of this is based on two studies performed for KMD by private market research companies, upon which the digitalization strategy relies heavily. The first, a quantitative study, found the following:

> A total of 58 per cent of the population and 64 per cent of business owners responded that they were satisfied with the public digital services. In the population, 30 per cent responded that they believed that digitalization should increase, 44 per cent that the current level was sufficient, and 13 per cent that digitalization had gone too far (KMD, 2019: 14).

These results are hardly a resounding assertion of discontent, with the majority expressing satisfaction with the status quo. Statistics such as the number of users of digital public services are included to justify demands. The other study was qualitative and found that the "user" wants "contact with the public sector that is fast, efficient and friction-free and that there is demand for, among other things, digital seamlessness between public organizations" (Kantar, 2019). The study questioned users regarding existing services and rule-based automation. Data-driven services were not explored. The study also found that few are concerned about the misuse of data by the public sector. This is relatively unsurprising, given the narrow scope of digitalization addressed. How, why, or on what premise the public sector should pursue datafication was not problematized.

We also discovered that many of the identified needs do not originate from direct engagement with citizens but from the public sector itself. Despite their connotations, user-centric needs-based approaches do not require citizen participation. The National Digitalization Council (2020) gives guidance that "one must stop to think about the user and start to think like the user." When asked how public entities approach datafication, we discovered that practitioners themselves are an important source of the "demanding user" who needs to be served. The actual end-user is often absent during the needs gathering phase. Instead, they are an imagined entity—user stories and customer journeys abound. One interviewee stated that "of course we all have user stories, and we follow them up in development and so forth" (Informant B). Another said, "This is really just kind of brainstorming. Coming up with lots of ideas, good ideas that are good for the citizens, or our customers in this case, then this is very good, I think so, at least." The interviewee went on to say, "This is what we mean, to give a better customer journey, better customer experience, user experience and all that" (Informant C). Several interviewees struggled to find demand for datafication from domain experts and street-level bureaucrats within their organization (Broomfield and Reutter, 2021). Needs then predominantly originated from data workers who placed themselves in the dual roles of end-user and supplier.

How, then, does the concept of user-centricity manifest itself? In short, user-centricity is about a solution rather than a process. Seamless services have been designated as the epitome of user-centric service delivery. The digitalization strategy states that user-centricity will be achieved through the development of seamless services based on life events, such as a "baby being born" (KMD, 2019: 16). Anticipatory services will be provided proactively, giving timely and personalized information and support while ensuring that data is easily accessible for all users. Everything is to be streamlined across organizational boundaries. Datafication is deemed a prerequisite for their realization: "Public sector shall exploit the potential of sharing and using data to create user-friendly services" (KMD, 2019: 4). Silos are to be torn down, allowing data to ebb and flow

across the public sector and, where relevant, also between the public and private sectors for co-creation of services.

We found examples of participation in seamless services. However, this is reserved for individual services. An example here is the "death and inheritance" service, which established 71 user needs from a broad mapping phase (Digdir, 2021). This mapping was conducted late in the process after many fundamental decisions had been made. A consultation was reserved for those who had been directly impacted by this particular life event. Citizens and civil society were not actively consulted on broader issues around datafication, such as where and how it should be applied, what should be prioritized, whether datafication is the optimal solution, and to what extent it should be implemented. Instead, those decisions were taken on their behalf.

Doing the best for society

Asserting that citizens are absent from datafication discourse in Norway would be an oversimplified account of the matter. As we have shown, citizens as both a discursive and actual participating entity are rarely engaged with in the early phases; however, they are considered. We frequently heard the phrase in interviews: "This should of course be done for the best of society." In this section, we dig deeper to ascertain what this might mean and how it manifests itself.

It was apparent during interviews that practitioners see their work as benefiting society. Much of this perspective translates to creating efficiencies and providing improved and seamless user experiences. One interviewee stated, "Everything we do is about being more effective. We must use less money and make lives easier for our customers" (Informant D). As one might expect, this mirrors what is mandated by policy. Few reflected critically beyond these issues. An exception here was a group of interviewees who raised the challenge of societal responsibility. One interviewee stated that it is very difficult to know how far to go because "Society needs to be in agreement with itself as to what it wants" (Informant E), to which a colleague agreed and elaborated, stating that, "This is way over our heads." As broader issues are neglected at the policy level, this pushes the responsibility for the interpretation of critical societal issues down to the organization and practitioner to interpret on a case-by-case basis.

Policy and practitioners also advocate for a variety of rights that need to be safeguarded in the recirculation of data, the guiding principle being that public bodies "shall share data when it can and protect data when it must" (KMD, 2019: 20). The idea of individual control and privacy is central here. Control issues are proposed to be solved by tools such as virtual assistants, which will give citizens insight into and control over their data (KMD, 2019: 16). Providing citizens with oversight on their data selves may contribute to legitimizing the recirculation of data and give a false sense of control. Non-sensitive data is also deemed unproblematic to recirculate and proclaimed to be in the best interest of the economy and society to share as open data. The potential for harm with this data is unexplored in the data. The lack of attention to this is of particular concern in the Norwegian data context, given the immense amounts of data that can be linked across registers. Even when anonymized, this can give deep insights into the population (Kitchin, 2014).

Policy and practitioners are not ignorant of challenges. Many of the submissions to the AI Strategy's consultation raised issues such as transparency, protection of public values, and the need to temper techno-optimism among politicians. The AI strategy regularly mentions ethics and refers to some challenges for the public sector, such as bias and explainability, but fails to further this discussion. Digdir and the Data Protection Authority are earmarked to provide guidance in this area, but this is still not forthcoming over 2 years after publication. Ethics scored low in the survey (Figure 1), with interviews exposing a general perception that if it's legal, it's ethical, and if it's legal, it's responsible. Many judged ethics to be irrelevant when working with non-sensitive data. Pertinent issues such as democratic oversight and the potential to change citizen–state relationships are not mentioned (Broomfield and Reutter, 2021) with no discussion around the potential for unforeseen consequences that many other jurisdictions have experienced (Alston, 2019; Gesley, 2020).

The digitalization strategy states that "The public sector shall be digitalized in a transparent, inclusive, and trustworthy way" (KMD, 2019: 8). Openness is regarded as important by practitioners, with one interviewee stating, "It's not like we're doing something all hush, hush in the back room. Everything is very transparent." The current openness regime is premised on citizens having both the competence and capacity to investigate and assess. The Norwegian population is believed to be digitally competent relative to other countries. This is measured by questions such as "Did you access service X online?" or "Have you opened a word file lately?" (Kompetanse Norge, 2018). Given the complexity involved in datafication, it is unlikely that such competence is a sufficient basis for non-expert citizens to scrutinize the public sector. Beyond encouraging uptake of the introductory "Elements of AI" course, we found no actions to deal with the new demands that datafication brings to the openness regime.

Discussion: conditions of participation in a datafied public administration

How are citizen perspectives problematized and included in policy and practitioner discourse in the datafication of public administration? Public sector datafication has the potential to substantially change power dynamics between citizen and state, as it enhances the state's ability to classify, control,

and predict citizen behavior, with the digital profiling of citizens emerging as a central action (Barassi, 2019; Hintz et al., 2019). Therefore, it is crucial to investigate the conditions of participation for citizens and civil society in its production and implementation. Applying Cardullo and Kitchin's (2019) framework to analyze overall discourse in the field, lifts the critique of datafied public administration to a system level and away from assessing single projects.

A recent study of the participatory turn in Norwegian public administration has found clear indications of citizen inclusion and grassroots involvement in climate change, energy, family, and gender policy fields (Krick et al., 2019). However, we find that this is not the case in the datafication domain. Citizens seem to be positioned within the "golden view" as passive stakeholders, often unable to engage with or challenge decisions that govern their lives (Dencik et al., 2019). This is despite great significance being placed on user-centric and needs-based approaches. If citizens are engaged, it is only at the latter stages, and if they belong to a specific user group, and they are not involved in many fundamental decisions. This would be akin to town planners developing a town plan together with builders and colleagues in the public sector and putting this forward to politicians without any public consultation around how the entire town will look. Possible alternatives for buildings and their use are never open for public debate but are pre-decided based on imagined user stories and builders' and architects' needs. Individuals are only invited into specific discussions on a particular building if they are deemed a potential "user" of that actual building. The architecture for Norway's datafied existence is decided on behalf of rather than together with society.

We found that civil society is rarely actively included by policy makers and practitioners regarding datafication, despite civil society being regarded as a vital actor in Norwegian democracy and considered crucial in policy production and the stimulation of debate during societal change (Meld. St. 10 (2018–2019)). In addition, the OECD (2017) has given clear recommendations to include civil society to bring different perspectives into the datafication discussion. The lack of involvement of civil society can perhaps be neatly explained away by the fact that there are few civil society actors in this field in Norway. There are, however, many international organizations, such as Access Now and Algorithm Watch, which could have been approached to provide a different perspective. There was no such hindrance to direct invitations for input to international companies such as Google, Microsoft, and IBM. This is an unexpected finding, as the Norwegian corporate model has a long tradition of including civil society and trade unions in policy and implementation processes (Christiansen et al., 2010).

Similar to van Zoonen's (2020) study in the Netherlands, we found datafication to be heavily influenced by internal public sector partnerships, leadership decisions, and close cooperation with the private sector, rather than by democratic processes and citizen participation. The private sector is actively involved in both design and implementation, regarded not only as a potential user or supplier of services but also as a partner and co-creator. This engagement is obvious in the resulting policy prioritizations. The private sector is a vital actor in the Norwegian corporatist state model; however, so too are other stakeholders. A variety of threats to the model have been identified, such as a more pluralist society, organized lobbyism, and increased expertization (Krick et al., 2019). Datafication is portrayed as a highly specialized topic with strong private-sector interests seeking to influence policy making. We found that only one hybrid committee has been proposed. There has also been minimal engagement with the broader academic community beyond the technical academic AI community.

Datafication, therefore, becomes a particular example of the erosion of the corporatist model. A debate seems unwelcome, with only cheerleaders for the positive imaginaries of datafication invited to dine at the table. As Mergel et al. (2018) point out, there is a substantial difference between commercial enterprises driven by a financial bottom line and the ambiguous and multifaceted public sector, which has the bottom line of creating public value. Cooperating predominantly with the private sector can limit what questions are raised, considered, and prioritized (Redden, 2018; Brauneis and Goodman, 2019). This further contributes to invoking an idea of a corporate/government inside with power and in control, and a disempowered and unknowing citizenry outside in the datafication of society. Datafication can lead to an increased blurring of the lines between the private and public sectors, even in countries such as Norway, where this divide has traditionally been strong (Crawford, 2021).

It could be argued that direct citizen involvement is unnecessary, as public servants can represent the citizen. They are, after all, 'users' of public services themselves and can, as the Digitalization Council recommends, "think like the user." However, this is fraught with difficulty. First, as Cardullo and Kitchin (2019) point out, we are seeing a proliferation of technical and management experts, such as data scientists, project managers, and architects. Social scientists and public administration experts, who may be better equipped to consider broader societal questions, are rarely included (Broomfield and Reutter, 2021). Second, datafication affects citizens differentially (Kennedy, 2018). Public servants are hardly representative of society at large and are unlikely to be the most impacted by datafication. Thirdly, the rhetoric around datafication may make it difficult for public servants to raise concerns. Fear of being dismissed as resistant to change, disloyal to the cause, or unpatriotic, as Sandvik (2020) experienced when questioning the Covid-19 app, are all factors that could hamper internal debate. We were unable to investigate whether there was internal resistance to policy development, as most internal policy production correspondence relevant for this paper falls outside of Freedom of Information legislation.

Taking Cardullo and Kitchin's (2019) framework into account, we find that the discourse in the field remains largely instrumental and paternalistic. This is surprising and a clear digression from the inclusive corporatist model, which despite coming under threat, is still loyally adhered to in many other domains of Norwegian policy making. Datafication policy is produced through a top-down approach of non-participation. During implementation, and particularly the design of services, participation may extend to consumerism and tokenism. We did not detect any attempt to facilitate citizen power. Citizens and civil society actors were virtually absent in the planning and decision-making processes.

The predominant user-centric approach seems little more than a re-branding of top-down technocratic efforts of datafication (Kitchin, 2014). The discursive turn from a citizen to a user is itself entangled with changing power relationships and fits into the overall paradigm of NPM and the reconstruction of citizens as consumers. It can be argued that the idea of the citizen as a user breaks with the disempowered welfare state client, as it attempts to strengthen citizen relations with the public sector and signals that people are given more influence over services (Langergaard, 2014; Sørensen, 2000). Conversely, it can be argued that users/customers have different rights and relations to the welfare state than citizens, as citizenship is associated with agency and responsibility (Mik-Meyer and Villardsen, 2012). The user is expected to accept asymmetrical power dynamics, often taking a passive role, dependent upon services, and given little agency other than how services might be designed (Gubrium and Järvinen, 2013). Blurring the lines between these concepts runs the risk of eroding the political and democratic conditions of public administration. This is further complicated by merging civil society, citizens, and the private sector into the single term of "user" in the Norwegian context (KMD, 2019). Early research on citizen participation predicted an increase in the direct involvement of citizens in democratic societies as they become more decentralized, interdependent, linked by ICTs, and challenged by 'wicked problems' (Roberts, 2004). This has, however, not transpired in the datafication of public administration in Norway. Imagining the user is not citizen involvement; it is neither participation nor co-creation. The idea of user-centered datafication remains largely tokenistic, with public administration owning and controlling all projects or co-creating them with the private sector.

Framed in overwhelmingly positive terms, datafication in Norway is regarded as necessary, inevitable, and steeped in opportunity. The analysis shows that technology and data are not regarded as objects entangled with power by practitioners and policymakers. Their political dimension is simply not acknowledged, echoing the academic fields' concern of presenting data-driven technology as neutral and apolitical entities (Crawford and Boyd, 2012; McQuillan, 2018). This becomes a materialization of dataism, which fosters a belief in data as the enabler of a better and more effective and objective society and, therefore, it is rarely questioned by involved parties (van Dijck, 2014).

Conclusion

This paper investigates how citizen perspectives are problematized and included in policy production and practitioner discourse in the datafication of public administration. It applies Cardullo and Kitchin's (2019) framework, extending it beyond smart cities to a system-level investigation of citizen participation. This serves to improve our understanding of how individuals are being turned into datafied citizens and the policies and political-economic structures that are making the datafication of the public sector possible (Barassi, 2019). We observe a paternalistic and top-down technocratic approach to citizen engagement in this administrative reform. Non-participation is particularly apparent at the policy level. We identify some tendencies toward consumerism and tokenism at the practitioner level; however, we fail to identify any evidence of citizen power. Bottom-up initiatives or grassroots contestations of datafication become difficult when the process itself is kept afar, and the discourse is paternalistic. Citizens and civil society are reduced to passive but demanding "users" to be served by the public sector. This is in direct contrast to the active engagement with the private sector during all phases—from policy production through to implementation.

This case study is limited in size; however, the system-level investigation highlights the importance of situated and context-specific approaches to public administration datafication. Further research is needed to deconstruct the inner workings of this administrative reform and the rationalities and processes behind it. Research should also investigate the intersection of the seemingly incompatible public administration paradigms of "datafication" and citizen participation.

The context, values, and agendas of datafication are often obscured from citizens. A crucial question is: Are citizens actively encouraged to discuss or challenge the datafication of public administration beyond how individual services might impact them? In short, the answer is "no." We find that they are not regarded as stakeholders or participants. Civil society is not invited to configure datafication beyond a few open calls for input. Citizens are neither included nor able to challenge the political rationalities shaping their datafied lives through institutionalized or more informal channels of participation. A public debate around datafication is deemed unnecessary in many Nordic countries (Snell and Tarkkala, 2019). Efforts seem to bypass democratic processes. The inclusive corporatist model so fundamental to Norwegian democracy is disintegrating in the datafication domain, as the socio-technical imaginary of data-driven public administration is presented as both inevitable and uncontestable.

Acknowledgments

All authors contributed equally to the research and publication of this paper. The authors would like to thank Alice Elizabeth Dalsjø for her research assistance in transcribing interviews and sorting literature. In addition, we would like to thank Malcom Langford and Emil Røyrvik for reading and commenting on earlier drafts of the article. We would also like to thank the Norwegian AI forum for its assistance in distributing the survey. We thank the three anonymous reviewers as well as the co-editor for their comments on this article

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Norwegian University of Science and Technology, Digital Transformation project, Digital Infrastructures and Citizen Empowerment, and Norges Forskningsråd Public Sector PhD scheme (grant number 299993).

ORCID iDs

Heather Broomfield https://orcid.org/0000-0001-6116-0794
Lisa Reutter https://orcid.org/0000-0002-4952-0416

Supplemental material

Supplemental material for this article is available online.

Note

1. The first analysis has resulted in a publication in the Scandinavian Journal of Public Administration. See Broomfield and Reutter (2021).

References

Alston P (2019) *Digital Technology, Social Protection and Human Rights. Report by the Special Rapporteur on Extreme Poverty and Human Rights*. A/74/493. New York, NY: United Nations. https://undocs.org/A/74/493.

Arnstein SR (1969) A ladder of citizen participation. *Journal of the American Institute of Planners* 35(4): 216–224.

Arter D (2004) Parliamentary democracy in Scandinavia. *Parliamentary Affairs* 57(3): 581–600.

Barassi V (2019) Datafied citizens in the age of coerced digital participation. *Sociological Research Online* 24(3): 414–429.

Benouaret K, Valliyur-Ramalingam R and Charoy F (2013) CrowdSC: Building smart cities with large-scale citizen participation. *IEEE Internet Computing* 17(6): 57–63.

Beraldo D and Milan S (2019) From data politics to the contentious politics of data. *Big Data & Society* 6(2): 1–11. doi:10.1177/2053951719885967

Berntzen L and Johannessen MR (2016) The role of citizen participation in municipal smart city projects: Lessons learned from Norway. In: Gil-Garcia JR, Pardo TA and Nam T (eds) *Smarter as the New Urban Agenda*. Cham: Springer, pp. 299–314.

Bingham LB, Nabatchi T and O'Leary R (2005) The new governance: Practices and processes for stakeholder and citizen participation in the work of government. *Public Administration Review* 65(5): 547–558.

Brauneis R and Goodman EP (2018) Algorithmic transparency for the smart city. *Yale JL & Tech* 20: 103.

Broomfield H and Reutter L (2021) Towards a data-driven public administration: An empirical analysis of nascent phase implementation. *Scandinavian Journal of Public Administration* 25(2): 73–97.

Cardullo P and Kitchin R (2019) Being a citizen in the smart city: Up and down the scaffold of smart citizen participation in Dublin, Ireland. *GeoJournal* 84(1): 1–13.

Chang L and Jacobson T (2010) Measuring participation as communicative action: A case study of citizen involvement in and assessment of a city's smoking cessation policy-making process. *Journal of Communication* 60(4): 660–679.

Christiansen PM, Nørgaard AS, Rommetvedt H, et al. (2010) Varieties of democracy: Interest groups and corporatist committees in Scandinavian policy making. *Voluntas: International Journal of Voluntary and Nonprofit Organizations* 21(1): 22–40.

Chun S, Shulman S, Sandoval R, et al. (2010) Government 2.0: Making connections between citizens, data and government. *Information Polity* 15(1, 2): 1–9.

Cornwall A (2002) *Locating citizen participation. IDS Bulletin* 33(2): 49–58.

Crawford K (2021) *Atlas of AI*. New Haven: Yale University Press.

Crawford K and Boyd D (2012) Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662–679.

Dencik L (2019) Situating practices in datafication – From above and below. In: Stephansen HC and Trere E (eds) *Citizen Media and Practices*. London and New York: Routledge, pp. 243–255.

Dencik L, Redden J, Hintz A, et al. (2019) The 'golden view': Data-driven governance in the scoring society. *Internet Policy Review* 8(2): 1–24.

Desrosières A (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge: Harvard University Press.

Difi-rapport (2018) Deling av data – konseptvalgutredning 2018: 7. Available at: https://www.digdir.no/digitalisering-og-samordning/deling-av-data-konseptvalgutredning/1412 (accessed April 2021).

Digdir (2021) Sluttrapport innsiktsfasen Livshendelsen dødsfall og arv. Available at: https://www.altinndigital.no/globalassets/doa/sluttrapport-dodsfall-og-arv-innsiktsfasen.pdf (accessed January 2021).

Ekspertgrupppen for Datadeling i Næringslivet (2020) Rapport fra ekspertgruppen for datadeling i næringslivet. Available at: https://www.regjeringen.no/contentassets/c98cce6745b0486c948c269dc80335c8/rapport-fra-datadelingsutvalget2.pdf (accessed September 2021).

Frank L (2000) When an entire country is a cohort. *Science (New York, N.Y.)* 287(5462): 2398–2399.

Gabrys J (2019) Data citizens – How to reinvent rights. In: Didier B, Isin E and Ruppert E (eds) *Data Politics: Worlds, Subjects, Rights*. London: Routledge, pp.248–266.

Gesley J (2020) *Netherlands: Court Prohibits Government's Use of AI Software to Detect Welfare Fraud*. Library of Congress.

Available at: https://www.loc.gov/law/foreign-news/article/netherlands-court-prohibits-governments-use-of-ai-software-to-detect-welfare-fraud/ (accessed 15 March 2020).

Gubrium JF and Järvinen M (2013) *Turning Troubles into Problems: Clientization in Human Services*. London: Routledge.

Hintz A, Dencik L and Wahl-Jorgensen K (2019) *Digital Citizenship in a Datafied Society*. Cambridge: Polity Press.

Hovde Lyngstad T and Skardhamar T (2011) Nordic register data and their untapped potential for criminological knowledge. *Crime and Justice* 40(1): 613–645.

Kantar. (2019) Rapport – Den digitale borger En kvalitativ studie av den digitale hverdagen – for Kommunal- og moderniseringsdepartementet. Available at: https://www.regjeringen.no/no/dokumenter/den-digitale-borger/id2637043/ (accessed 6 October 2021).

Kaun A and Dencik L (2020) Datafication and the welfare state. *Global Perspectives* 1(1): 1–8.

Kennedy H (2018) Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication. *Krisis: Journal for Contemporary Philosophy* 2018(1): 18–30.

Kennedy H, Oman S, Taylor M, et al. (2020) Public understanding and perceptions of data practices: A review of existing research. Living with Data. Available at: https://livingwithdata.org/current-research/publications/ (accessed 10 March 2021).

Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London: Sage Publications.

Kitchin R (2017) Thinking critically about and researching algorithms. *Information, Communication & Society* 20(1): 14–29.

Kluge Advokatfirma AS (2020) *Utredning – Rettslig Handlingsrom for et Generelt Pålegg mot Offentlige Organer og Virksomheter om Aktivt å Tilgjengeliggjøre Offentlig Informasjon*. Ministry of Local Government and Modernisation. Oslo. Available at: https://www.regjeringen.no/no/dokumenter/utredning–rettslig-handlingsrom-for-et-generelt-palegg-mot-offentlige-organer-og-virksomheter-om-aktivt-a-tilgjengeliggjore-offentlig-informasjon/id2722129/ (accessed 10 March 2021).

KMD (2019) En digital offentlig sektor: Digitaliseringsstrategi for offentlig sektor 2019–2025. Available at: https://www.regjeringen.no/no/dokumenter/en-digital-offentlig-sektor/id2653874/ (accessed 1 December 2019).

KMD (2020) Nasjonal strategi for kunstig intelligens. Available at: https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/ (accessed 3 February 2021).

KMD (2021) Melding om datadrevet økonomi og innovasjon. Available at: https://www.regjeringen.no/no/tema/statlig-forvaltning/ikt-politikk/melding-om-datadrevet-okonomi-og-innovasjon/id2688376/ (accessed 26 March 2021).

Kompetanse Norge (2018) Grunnleggende digitale ferdigheter i befolkningen. Available at: https://www.kompetansenorge.no/statistikk-og-analyse/grunnleggende-digital-ferdigheter (accessed 1 September 2019).

Krick E, Christensen J and Holst C (2019) Between 'scientization' and a 'participatory turn'. Tracing shifts in the governance of policy advice. *Science and Public Policy* 46(6): 927–939.

Krick E and Holst C (2021) Governance by hybrid advisory committees – a hallmark of social democracy? In: Bratberg Ø and Thorsen DE (eds) *Social Democracy in the 21st Century (Comparative Social Research*, Vol. 35*)*. Bingley: Emerald Publishing, pp.113–130.

Langergaard LL (2014) 4 User-Driven innovation and the role of the citizen. In: Fuglsang L, Rønning R and Enquist B (eds) *Framing Innovation in Public Service Sectors*. New York: Routledge, pp.63–83.

Lehtiniemi T and Ruckenstein M (2019) The social imaginaries of data activism. *Big Data & Society* 6(1): 1–12. doi:10.1177/2053951718821146

Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.

McQuillan D (2018) Data science as machinic neoplatonism. *Philosophy & Technology* 31(2): 253–272.

Meld. St 10 (2018–2019) (2019) *Frivilligheita – sterk, sjølvstendig, mangfaldig — Den statlege frivilligheitspolitikken*. Oslo: Ministry of Culture.

Mellouli S, Luna-Reyes LF and Zhang J (2014) Smart government, citizen participation and open data. *Information Polity*, 19(1, 2): 1–4.

Meng A and DiSalvo C (2018) Grassroots resource mobilization through counter-data action. *Big Data & Society* 5(2): 1–12. doi:10.1177/2053951718796862

Menon Economics (2019) Er verdiskaping med data noe Norge kan leve av? På oppdrag fra NHO. Available at: https://www.menon.no/publication/verdiskaping-data-norge-leve/

Mergel I, Rethemeyer RK and Isett K (2016) Big data in public affairs. *Public Administration Review* 76(6): 928–937.

Mik-Meyer N and Villardsen K (2012) *Power and Welfare: Understanding Citizens' Encounters with State Welfare*. London: Routledge.

Milan S (2016) Data activism as the new frontier of media activism. In: Yang G and Pickard V (eds) *Media Activism in the Digital Age*. London: Routledge, pp. 151–163.

Misuraca G and van Noordt C (2020) *Artificial Intelligence in Public Services - Overview of the use and Impact of AI in Public Services in the EU*. EUR 30255 EN. Luxembourg: Publications Office of the European Union.

Mossberger K, Tolbert CJ and McNeal RS (2007) *Digital Citizenship: The Internet, Society, and Participation*. Cambridge, MA: MIT Press.

NOU 2003:19 (2003) *Makt og demokrati: Sluttrapport fra Makt- og demokratiutredningen*. Statens forvaltningstjeneste Informasjonsforvaltning Oslo. Available at: https://www.regjeringen.no/contentassets/316f4765f7c44a2c8def9dcdb5da8f30/no/pdfs/nou200320030019000dddpdfs.pdf. (accessed 10 March 2021).

OECD (2017) Digital government review of Norway - boosting the digital transformation of the public sector. *OECD Digital Government Studies*. Available at: https://www.oecd.org/gov/digital-government/digital-government-review-norway-recommendations.pdf (accessed 10 March 2021).

Ramaswamy V and Ozcan K (2014) *The Co-Creation Paradigm*. Stanford, California: Stanford University Press.

Redden J (2018) Democratic governance in an age of datafication: Lessons from mapping government discourses and practices. *Big Data & Society* 5(2): 1–13. doi:10.1177/2053951718809145

Roberts N (2004) Public deliberation in an age of direct citizen participation. *The American Review of Public Administration* 34(4): 315–353.

Rokkan S (1966) *Norway: Numerical Democracy and Corporate Pluralism*. Bergen: Chr. Michelsens institutt.

Ruppert E, Isin E and Bigo D (2017) Data politics. *Big Data & Society* 4(2): 1–7. doi:10.1177/2053951717717749

Sandvik KB (2020) "Smittestopp": if you want your freedom back, download now. *Big Data & Society* 7(2): 1–11. doi:10.1177/2053951720939985

Snell K and Tarkkala H (2019) Questioning the rhetoric of a 'willing population' in Finnish biobanking. *Life Sciences, Society and Policy* 15(1): 4.

Sørensen E (2000) Democratic governance and the changing role of users of public services. *Administrative Theory & Praxis* 22(1): 24–44.

Torfing J, Sørensen E and Røiseland A (2019) Transforming the public sector into an arena for co-creation: Barriers, drivers, benefits, and ways forward. *Administration & Society* 51(5): 795–825.

Tupasela A, Snell K and Tarkkala H (2020) The nordic data imaginary. *Big Data & Society* 7(1): 1–13. doi:10.1177/2053951720907107

van Dijk J (2014) Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society* 12(2): 197–208.

van Ooijen C, Ubaldi B and Welby B (2019) A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance. In *OECD Working Papers on Public Governance*. Paris: OECD Publishing. at: https://www.oecd-ilibrary.org/governance/a-data-driven-public-sector_09ab162c-en (accessed 25 January 2020).

van Zoonen L (2020) Data governance and citizen participation in the digital welfare state. *Data & Policy* 2. doi:10.1017/dap.2020.10

Yeung K (2017) 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society* 20(1): 118–136.

Yeung K (2020) From new public management to new public analytics. In: ADM workshop 1: conceptualizing automation in the welfare sector, 16 November 2020.

Zuboff S (2019) *The Age of Surveillance Capitalism*. New York: Profile Books.

[The Supplementary material is available on the Big Data & Society Web Page. It is here attached for the consistency of cross references]

Appendix A: Overview over key policy documents in the field of public administration datafication in Norway

(Paper 3)

Title and date	Produced by/Ordering entity	Content (as described by document)
Digital Agenda for Norway White Paper, Meld. St. 27 (2015–2016)	Ministry of Local Government and Modernisation	The report presents the government's main goals and priorities in ICT policy for a user centric and efficient public administration, value creation and inclusion. It describes how ICT can be used to renew, simplify and improve the public sector. It further discusses how ICT can facilitate innovation and competitiveness in business and industry and presents a national plan for electronic communication.
Concept Phase Study for sharing of Data November 2018	Norwegian Digitalization Agency	The goal of the study was to identify concepts which would make it easier for the public and private sector to access and use public data. This would in turn realise a more efficient and user oriented public sector and value creation in the private sector. The report recommends an agile approach where both the public and private sectors think big, start small and learn by doing.
Digitalization strategy for the public sector "One digital public sector", May 2019	Ministry of Local Government and Modernisation	The strategy defines the common goals and focus areas for digitalisation activities towards 2025. Digitalisation of the public sector aims to give citizens, businesses, and the voluntary sector a simpler everyday life through better services and more efficient use of resources by government agencies, and facilitate increased productivity in society at large. The purpose of the strategy is to support a digital transformation in the individual agencies and in the public sector as a whole.

National Artificial Intelligence strategy for Norway, January 2020	Ministry of Local Government and Modernisation	The National Strategy for Artificial Intelligence is intended for the civilian sector – both private and public. It does not cover the defence sector. The strategy focuses on specifying what is meant by artificial intelligence and on describing some areas where it will be important for Norway to exploit the opportunities offered by AI.
White Paper for Data Driven Economy and Innovation 2021	Ministry of Local Government and Modernisation	The purpose of the white paper is to lay out the Governments policy for value creation with data as a resource. This is based on the premise that the data economy is an important driver for economic growth for Norway. The government wants Norway to take advantage of the opportunities offered by data for increased value creation, new jobs throughout the country, and an efficient public sector. Better utilization of data is important if Norway is to succeed in the transition to a more sustainable society and a greener economy. The Government's ambition is to increase the sharing of data within the business community and between the public and private sectors.

Appendix B: Survey template (Paper 3)

1) In which organisation are you employed? [Free text question]
2) What job code/title do you have? [Free text question]
3) In your opinion, how data-driven is your organisation in relation to other actors in the public sector? Please answer this on a scale of 1 to 10, with 10 being very data driven and 1 being minimally data driven. [Likert scale]
4) In your opinion, how highly prioritised is AI/data science in your organisation? Please answer this on a scale of 1 to 10, with 10 being highly prioritised and 1 being not prioritised. [Likert scale]
5) How is the work with AI/data science organised within your organisation? Here, examples of answers are (not exhaustive): the IT department; the statistics/analysis department; integrated across the organisation. [Free text question]

6) Who are the main AI model developers for your organisation? [Drop-down list]
    a) Mainly in-house staff
    b) Approximately 50/50 consultants and staff
    c) Mainly hired consultants
    d) Other
7) How far has your organisation progressed in the work with AI/data science? [Drop-down list]
    a) Starting to think about using it
    b) Planning phase
    c) Start-up phase
    d) Testing
    e) Production
8) Operation/management
9) Follow up to 7: (a) What are you considering using AI/data science for? [Free text question]; (b) What are you planning to use AI/data science for? [Free text question]; (c) What is AI/data science used for in your organisation? [Free text question]
10) Below are a number of potential reasons for using AI/data science. Please rank them from most important (highest) to least important (bottom). [Ranking]
    a) More efficient decision-making processes
    b) Better quality and timeliness in decisions
    c) More precise predictions
    d) Increased user orientation
    e) Increased innovation and business development
    f) Increased employee satisfaction
    g) Reduced costs

11) Below are a number of potential challenges around the adoption of AI in the public sector. Please rate how large the challenge is for your organisation on a scale from 1 to 5, with 1 being no challenge and 5 being a very large challenge. [Likert scale]
    a) Technical infrastructure
    b) Ethics
    c) Organisational culture
    d) Access to data
    e) Data quality
    f) Privacy and security
    g) Analytics competence
    h) Legal and regulatory framework
    i) Citizens' insecurity and willingness to accept AI/data use
    j) Uncertainty around what AI could be used for in the organisation
    k) Lack of funding
    l) Pressure from management to deliver
    m) Fear of downsizing in the organisation
12) Are there other challenges when using AI/data science in your organisation? [Free text question]
13) Do you have a project/activity in the AI/data science area that you would like to tell us about so that we can share it with others? Please give a short description of the project/activity. [Free text question]
14) Would you or anyone else in your organisation be willing to be interviewed by us? We are looking for insights into what the public sector needs in the AI/data science field. [Yes/No]
    a) Contact information [Free text question]
15) Do you have anything else you would like to add or comments on this questionnaire? [Free text question]

Appendix C: Interview guide (Paper 3)

The guide followed a general approach, with a number of common questions posed to each interviewee based on the survey questions, which are found below. An individual guide was, however, prepared for each organisation, adapted to the responses that were made in the survey. Here, we made comments on interesting issues to follow up with the interviewee(s). We presented the responses to the interviewees and encouraged them to elaborate and explain their justifications for the responses.

Questions

1. What is your name and job title, and what are you working with in your organisation?
2. In your opinion, how data-driven is your organisation in relation to other actors in the public sector?
    a. Why did you position your organisation here?
    b. What does it mean to your organisation to become data driven?
3. In your opinion, how highly prioritised is AI/data science in your organisation?
4. How is the work with AI/data science organised within your organisation?
    a. Why is the work organised this way?
5. Who are the main AI model developers for your organisation?
    a. Why did the organisation choose to place the responsibility here?
6. How far has your organisation progressed in the work with AI/data science?
    a. What are you currently planning to use this technology for, or what is it already used for?
    b. Do you have a project/activity in the AI/data science area that you would like to tell us about?
7. In the survey, you ranked potential reasons for using AI/data science in your organisation. Can you elaborate on this ranking?
    a. Are there any other reasons that you would like to add?
8. Graph: Here we presented the interviewee(s) with the graph of their results measured against the average of the AI forum and encouraged them to discuss each of the challenges and justifications as to their ranking. They were asked to elaborate on each of the challenges.
    a. Are there other challenges when using AI/data science in your organisation?
    b. Would you like to discuss challenges in a specific project?
9. Is there anything else you want to discuss or bring up?

[Paper 4]
Public Sector Data as a Resource: Tracing the Emergence and Embedding of a Sociotechnical Imaginary.

This paper is co-authored with Heidrun Åm and currently under process at *Critical Policy Studies*

Decision in first round: Revise and Resubmit

Appendix A : Norwegian Centre for Research Data

Assessment of research project. First assessment was given 1 February 2019. The project period was then extended 3 August 2021. Two assessments are therefore attached.

Vurdering

Dato
01.02.2019

Type
Standard

Referansenummer
780073

Prosjekttittel
(Re)constructing algorithmic governance in the Norwegian public sector-between citizen empowerment and technocratic rationality

Behandlingsansvarlig institusjon
Norges teknisk-naturvitenskapelige universitet / Fakultet for samfunns- og utdanningsvitenskap (SU) / Institutt for sosiologi og statsvitenskap

Prosjektansvarlig
Lisa Reutter

Prosjektperiode
01.08.2018 - 01.08.2021

Meldeskjema

Kommentar
Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 01.02.2019. Behandlingen kan starte.

MELD ENDRINGER
Dersom behandlingen av personopplysninger endrer seg, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. På våre nettsider informerer vi om hvilke endringer som må meldes. Vent på svar før endringer gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 01.08.2021.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).
NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.
Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).
Dersom du benytter en databehandler i prosjektet må behandlingen oppfylle kravene til bruk av databehandler, jf. art 28 og 29.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig

institusjon.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp underveis og ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!
Tlf. Personverntjenester: 55 58 21 17 (tast 1)

Vurdering

Dato
03.08.2021

Type
Standard

Referansenummer
780073

Prosjekttittel
(Re)constructing algorithmic governance in the Norwegian public sector-between citizen empowerment and technocratic rationality

Behandlingsansvarlig institusjon
Norges teknisk-naturvitenskapelige universitet / Fakultet for samfunns- og utdanningsvitenskap (SU) / Institutt for sosiologi og statsvitenskap

Prosjektansvarlig
Lisa Reutter

Prosjektperiode
01.08.2018 - 01.11.2022

Meldeskjema

Kommentar
NSD har vurdert endringen registrert 2.8.2021.

Endringen innebærer at prosjektslutt utsettes til 1.11.2022 (tidligere 1.8.2021).

Det er vår vurdering at behandlingen fortsatt vil være i samsvar med personvernlovgivningen, så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjema med vedlegg 3.8.2021. Behandlingen kan fortsette.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Kontaktperson hos NSD: Lasse Raa

Lykke til videre med prosjektet!

Appendix B: Public Dissemination

Presentations/public talks (not including guest lectures)

Title (date)	Event/Audience	Content
Bransjen vs. Akademia: Et første forsøk på å bygge bro (2019)	Make Data Smart Again 2019. Conference for data scientists and others working on data-driven technology across the public and private sector, organized by the Norwegian Data Association (Den norske dataforeningen)	Keynote where I presented a first literature review on how social scientists perceive the data-driven paradigm, focusing on power imbalances and AI hype.
AI + offentlig sektor =sant? (2019)	AI-konferansen. Conference organized by the Norwegian Data Association (Den norske dataforeningen). Audience mostly made up of data scientists and other IT staff across the public and private sector.	Presentation on the first findings of the analysis presented in paper 2.
Et kunstig intelligent samfunn – Hvordan skal din digitale fremtid se ut? (2020/2021)	Researcher's night. Audience made up of high-school students.	Presentation about AI in society and its shortcomings. Using examples to show how technology is never perfect. Thought experiment on predicting school dropouts.
Overlat byråkratiet til robotene (2020)	Podcast, Uillustrert vitenskap, produced by Radio Revolt, popular science for students.	Discussion about the use of automated decision-making in public administration.
Digitalisering og datadrevet offentlig forvaltning fra et samfunnsvitenskapelig ståsted (2021)	KAI konferansen 2021. Conference organized by the communal archival institutions. Audience mainly made up of public sector employees working with archival tasks.	Presentation about data-driven public administration. Presentation of findings from paper 2 and social science research on datafication more generally.
Lærende maskiner, klokere forvaltning? Om bruk og implementering av maskinlæring i offentlig sektor (2021)	eForvaltningskonferansen 2021 Conference on eGovernment organized by several trade unions. Audience mainly made up of public sector employees.	Presentation about data-driven public administration. Presentation of findings from paper 2 and social science research on datafication more generally.

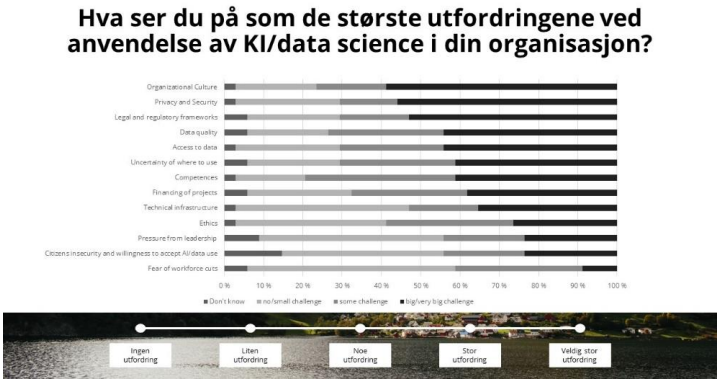
Title (date)	Event/Audience	Content
Kunstig intelligens – status, muligheter og utfordringer (2021)	Meeting of the IT-leader network (SnIT) organized by the Norwegian Data Association (Den norske dataforeningen). Audience made up by IT leaders across the private and public sector.	Presentation about social science perspectives on data-driven technology, focusing on findings of papers 1 and 2 and unseen issues.
Demokratisk digitalisering – Hvem skal styre våre data (2021)	Panel organized by ATTAC Norway. Audience mostly made up of students interested in data politics.	Presentation introducing research on the datafication of society. Short introduction to my own work in the field. Panel discussion on data politics.
Et datadrevet samfunn (2021)	NOKIOS. Workshop 2 : Surveillance Economy Out of Control—Commercial Tracing in the Public Sector. NOKIOS is a conference organized by the Norwegian University of Science and Technology and actors from the public sector, bringing together both private and public sector organizations working in ICT in the public sector.	Presentation on data-driven public administration and critical perspectives on data sharing and recirculation in the public sector. Followed by a panel discussion on the use of commercial platforms in public administration.
Ingenting å skjule: registerovervåling (2022)	Podcast produced by the Norwegian Data Protection Office	Interview about the role of register data in the welfare state and how this changes when the idea of data-driven public administration is introduced.
Om å bli datadrevet – store og små utfordringer for offentlig forvaltning (2022)	eKommune 2022. Conference organized by the Norwegian Association of Local and Regional Authorities. Audience made up of public sector employees in local governments in Norway.	Presentation about what data-driven public administration is and its practical and societal challenges. Critical perspective on public administration datafication and its framing in Norway. Obstacles as moments of reflection
Datadrevet offentlig forvaltning: En kort innføring i kunstig intelligens og dens begrensninger (2022)	Meeting of the Opptaksutvalget, an official public committee commissioned by the Ministry of Education and Research to assess future practices of higher education admission.	Presentation about data-driven public administration, AI and automatic decision making in higher education admission.

Written dissemination

Autor (date) Title	Genre	Content
Reutter, L. (2019) Kun en av fem offentlige virksomheter bruker kunstig intelligens – og det er kanskje bra.	Opinion piece in digi.no	This is answering the administrative director of ICT-Norway, who stressed that only one out of five public organizations makes use of AI. According to her, this is highly problematic. In this opinion piece, I argue that it might not be that bad that not every organization is using AI as there are various unresolved issues tied to implementing data-driven technology in the public sector.
Reutter, L. and Broomfield, H. (2019) Kunstig intelligens/data science: En kartlegging av status, utfordringer og behov i norsk offentlig sektor – en første analyse.	Report to the Ministry of Modernization and Local Government	Report presenting the first findings of survey and interviews (initial analysis of paper 1) sent to the Ministry as input to the work on the national AI strategy.
Norges forskningsråd (2020) Hvordan kan en chatbot bli rassist over natt?	Expert interview published in forskning.no	Interview on the social impact of data-driven technology
Reutter, L. and Grolid S. Å. (2020) Kunstig intelligens	Digital teaching resources for high school teachers, Nasjonal Digital Læringsarena (The Norwegian National Digital Learning Arena)	Developed digital learning resources on the relationship between AI and society for high school teaching. Included written information and illustrations about how AI works and a thought experiment about the use of data-driven technology in predicting school drop-outs.

Autor (date) Title	Genre	Content
Aasback, A. W. and Reutter, L. (2021) Kunstig intelligens i sosialt arbeid	Article in Fontene, a journal published by Fellesorganisasjonen (trade union for social workers)	Introduction to AI and automated decision-making for social workers. Critical reflections on the role of subject matter expertise in producing data-driven technology.
Tønnesen, H.; Reutter, L. and Magin, M. (2021) Når alt vi gjør registreres og systematiseres: Stordata i hverdagen	Chapter in a popular science book on the digital everyday in Norway. Edited by Rolstadås, Krokan, Dahle Øien, Rolfsen, Sand, Syse, Husby, and Waag	Introduction to big data and its role in society
Stølen, H. (2021) Tiktok anklages for sensur	News article on nrk.no. Expert interview.	Interview about automatic content moderation on internet platforms.
Gundersen, M. (2022) SSB krever å få vite nøyaktig hva nordmenn kjøper i matbutikken	News article on nrkbeta.no. Expert interview.	Interview about the use of transactional data by Statistics Norway. Contextualizing new data practices in the new data-driven paradigm.

Appendix C: Example Presentation

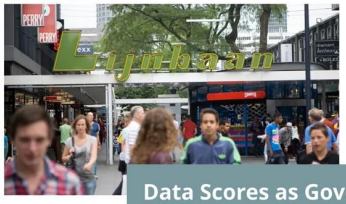


Datadrevet offentlig forvaltning

Mer, annerledes
og ny data

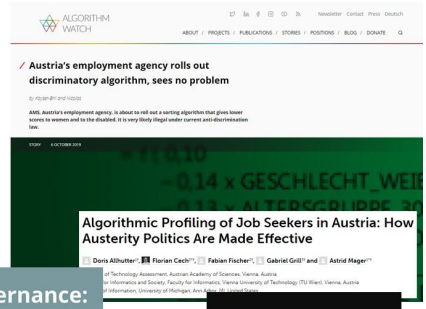
✚

Avanserte metoder
(maskinlæring, kunstig intelligens)
for å analysere data og resirkulere
data i praksis



Welfare surveillance system violates human rights, Dutch court rules

Government told to halt use of AI to detect fraud in decision hailed by privacy campaigners



Austria's employment agency rolls out discriminatory algorithm, sees no problem

Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective

Data Scores as Governance: Investigating uses of citizen scoring in public services

Project Report

Lina Dencik, Arne Hintz, Joanna Redden & Harry Warne

Hva snakker vi lite om i diskusjonen om data-drevet offentlig sektor?

Offentlig sektor er annerledes enn privat sektor

'Many governmental decisions are especially weighty and democratically elected governments have special duties of accountability' (Brauneis & Goodman 2018)

'Data-driven systems might transform state-citizen relations and understandings of both people and social issues.' (Dencik et. al 2019)

'For commercial enterprises, the exploitation of these data can often turbocharge the financial bottom line. However, for public organizations the ambiguous, multifaceted, and **contested "bottom line" of creating "public value" generates a set of important questions and concerns.**' (Mergel et. al)

ISBN 978-82-326-5905-0 (printed ver.)
ISBN 978-82-326-6686-7 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology