BECOMING DATA-DRIVEN?

The reconfiguration of work situations in the Danish Customs and Tax Administration

Bastian Jørgensen

IT University of Copenhagen | PhD Dissertation

BECOMING DATA-DRIVEN?

The reconfiguration of work situations in
the Danish Customs and Tax Administration

A dissertation submitted in compliance
with the requirements for the degree of

DOCTOR OF PHILOSOPHY (PH.D.)

Bastian Mosbæk Jørgensen
bmjo@itu.dk
Department of Business IT

IT University of Copenhagen
Submitted 29 March 2021

Supervisor
Christopher Gad
Associate Professor
IT University of Copenhagen

Co-supervisor
Brit Ross Winthereik
Professor
IT University of Copenhagen

IT-UNIVERSITETET I KØBENHAVN

ABSTRACT

This dissertation explores how work situations become reconfigured as public organizations strive to become data-driven. In 2016, the Danish Government published a new digital strategy encouraging public organizations to take the next step in modernization. In line with the ongoing popular rhetoric of an unfolding data revolution, the digital strategy emphasized data as an essential asset of the future digital age. Public organizations were recommended to invest in, develop, and improve upon their data infrastructures. The digital strategy argued that by increasing their use of data, public organizations would become more efficient and able to deliver better public services to citizens. The strategy is one key source for the emergence of a new and popular vision inside public organizations: To become data-driven. This dissertation studies the emergence of this data-driven vision in the Danish Customs and Tax Administration. Using an ethnographic approach, the dissertation explores how data-driven visions and technologies affect public organizations' work and organization. It describes and analyzes how organizational relations and boundaries become reconfigured, as public employees work with emerging data infrastructures. The study draws on science and technology studies, critical data studies, and public administration theory. The dissertation consists of three research papers, each analyzing a work situation affected by data-driven visions and technologies. In the first work situation, we are introduced to frontline workers who are confronted with the idea of making taxpayers digital and self-serving, also referred to as 'no-touch customers'. In the second work situation, we are introduced to data scientists and customs officers experimenting with a new machine learning algorithm for fraud detection. Finally, in the third work situation, we follow the project participants of a large IT project as they attempt to 'bring home' control and responsibility of data sharing in the public sector. In each of these work situations, a central figure appeared, which had the function of connecting the task of implementing new information technologies and visions to existing organizational challenges and decisions. The dissertation traces the different figurations and their practical effects, highlighting the emergence of new organizational boundaries, including how organizational competencies and responsibilities become reallocated between public employees, and the shifting relations between what counts as visible and invisible work. The dissertation contributes with new knowledge about how a public organization and its employees are affected by data-driven visions and technologies.

RESUMÉ

Denne afhandling undersøger hvordan arbejdssituationer bliver rekonfigureret som følge af offentlige organisationers stræben efter at blive data-drevne. I 2016 udgav Regeringen sammen med KL og de Danske Regioner en ny digitaliseringsstrategi *Et stærkere og mere trygt digitalt samfund*. I strategien blev offentlige organisationer anbefalet at tage det næste skridt i den digitale omstilling. Strategien lagde sig op af en populær retorik om en igangværende data revolution og fremhævede data som et helt centralt råstof i den digitale tidsalder. Offentlige organisationer blev anbefalet at investere i og videreudvikle deres data infrastrukturer for at øge effektiviteten og forbedre det offentliges service til borgere. I de senere år har flere offentlige organisationer påbegyndt arbejdet mod at blive mere data-drevne. Denne afhandling undersøger hvordan visionen om at blive data-drevet er med til at forandre arbejdet, organiseringen, og brugen af teknologi i offentlige organisationer. Afhandlingen tager udgangspunkt i en række etnografiske nedslag fra den danske skatteforvaltning. Den undersøger hvordan data-drevne visioner og teknologier er med til at forandre arbejdssituationer i det danske skattevæsen. Hvilke nye forventninger opstår til de offentlige ansatte, og hvilket arbejde må de udføre for at tilpasse nye teknologier og visioner til eksisterende arbejdssituationer? Afhandlingen undersøger disse spørgsmål ved primært at trække på forskningsfeltet videnskabs- og teknologistudier og endvidere på kritiske data studier og studier indenfor digitaliseringen af den offentlige administration. Afhandlingen består af 3 forskningspublikationer, der beskriver og analyserer hvordan arbejdssituationer i skatteforvaltningen bliver påvirket af data-drevne visioner og teknologier. Den første arbejdssituation analyserer hvordan kommunikationen mellem frontmedarbejdere og borgere er blevet påvirket af nye teknologier og visionen om den digitale og selvforsynende borgere. Den anden arbejdssituation omhandler arbejdet med at udvikle en maskinlæringsalgoritme der skal gøre toldkontrollen i skatteforvaltningen mere effektiv og præcis. Til sidst, i den tredje arbejdssituation, følger vi en gruppe projektdeltageres forsøg på at 'hjemtage' kontrollen og ansvaret for udviklingen af informations teknologier og delingen af skattedata med eksterne organisationer. I hver af disse arbejdssituationer fremhæves en central figure som forbandt implementeringen af de nye data-drevne visioner og teknologier med eksisterende organisatoriske problemstillinger. Afhandlingen følger disse figurer og deres effekter i de konkrete arbejdssituationer, og beskriver fremkomsten af nye organisatoriske grænser. Afhandlingen sætter især fokus på hvordan ansvar og kompetencer bliver omfordelt mellem offentligt ansatte, borgere og eksterne aktører og hvilket arbejde der bliver fremhævet og værdsat og dertil hvilket arbejde der bliver mere usynliggjort. Afhandlingen bidrager med ny viden om hvordan arbejdssituationer i offentlige organisationer bliver forandret som følge af data-drevne visioner og teknologier.

ACKNOWLEDGEMENTS

Handing in this dissertation marks the end of a long journey, and it is no secret that this final document only represents a fraction of my experiences as a PhD student. I am really grateful for having had the opportunity to pursue this PhD and especially for meeting so many kind, generous, and knowledgeable people along the way. So many people have a role in making this dissertation come through, and I owe a large thanks to everyone who has supported me, your help has been indispensable.

A special thanks to my supervisor Christopher Gad, who has supported me through thick and thin, always reminding me, in his calm and patient fashion, that doing research takes time. Thanks, Christopher, for always having your door open for numerous conversations, and for being patient with me in times of my restlessness and indecisiveness. Another tremendous thanks to my co-supervisor Brit Ross Winthereik. Thanks for letting me in on the Data as Relation research project, for always believing in me, and for pushing me to do my very best. You have done a remarkable job in creating a research environment that has been inspiring, open, speculative, daring, crooked, intellectually challenging, fun, and just wonderful to be part of.

Being part of the Data as Relation research project, has shaped my experience as a PhD student in many positive ways. Thanks to the fellow PhD students in the project, and good friends, John Burnett, Michael Hockenhull, and Jannick Schou, whom it has been a pleasure to share office with for a large part of the time. I apologize for being chatty at times, although I want to add that it has almost always been research related. A thanks to everyone else in the Data as Relation research project: James Maguire, Rachel Douglas, Marisa Cohn, Irina Papazu, Caroline Salling, Lise Røjskjær Pedersen, Morten Hjelholt, Mace Ojala, Lea Enslev, Luca Rossi, and Christina Neumayer. Thanks to ETHOS Lab manager Marie Blønd, and thanks to Ester Fritsch for the companionship along the way, including throughout several PhD courses, and in Bæredygtig Digitalisering Lab with Michael and John.

As a PhD student, I have been lucky to be part of many different research communities. Early on during my research, I was in a confused state about belonging to a specific research community. In this moment of despair, Helen Verran 'knighted' me a Science and Technology Studies (STS) scholar. Whatever the fluid identity of an STS scholar entails, I do think about the following people as part of my STS community. Thanks to the Data as Relation advisory board for their comments and inspiration: Geoffrey Bowker, Adrian McKenzie, Evelyn Ruppert, Casper Bruun Jensen, and Mark Elam. Thanks to my mid-way panel: Peter Danholt,

Karen Boll, and Baki Cakici. Thanks also to the research department of Organization, Work, and Technology at Lancaster University, and especially to Theodore Vurdubakis for hosting me at Lancaster University and for taking the time to listen to my research, provide valuable feedback, and for checking-in from time to time.

I could not have met all these people if it had not been for the funding of the Velux Foundation and the administration and support from the PhD School. Thanks to Vibe Mathiasen, Julie Jacobsen, and the head of the PhD School, Sisse Finken, for all the support, and for making it possible to hand-in this dissertation, especially when pressed for time. And a large thanks to Elisabeth Paul for thoroughly editing and commenting on the dissertation on such short notice.

I also want to express my gratitude towards the many employees in the Danish Customs and Tax Administration, who have so generously set aside time to introduce me to their work. It is no small thing to include a researcher in an already hectic workday. It has been extremely interesting for me to hear many diverse perspectives on how digitalization affects a public organization, and I hope that I have managed to faithfully include these diverse perspectives in the present dissertation.

A final thanks goes out to my family. In my work, I have been inspired by the curiosity and straightforwardness of my father, and the critical thinking and outspokenness of my mother – thanks to both of them for their love and support. Also, thanks to my parent-in-laws for lending me their house for writing, and for supporting me, my girlfriend Josefine, and our daughter Karla, especially when the Corona virus made its entry. Last, and most importantly, I owe more than words can express to Josefine for enduring this crazy process with me. While it is hard work to do a PhD, it might be even harder to stand on the side being supportive throughout the constant ups and downs while taking on extra work at home. It is needless to say that without the love and support from Josefine I could not have written this dissertation. Thanks Jose for always believing in me and for pushing me to reach the finish line. The joy and laughter we have shared together with Karla has been crucial for providing me with the energy to keep this project afloat.

Bastian Jørgensen
Copenhagen, 25th of March 2021.

TABLE OF CONTENTS

PART II

PART I

1 INTRODUCTION

The first chapter is an introduction to the thesis. It begins with an ethnographic montage that takes the reader stepwise closer to the Danish Customs and Tax Administration's (DCTA) endeavors to become data-driven. The montage reflects my intent with the dissertation, which is to situate the vision of becoming data-driven and its effects in work situations in the DCTA. It highlights how new data technologies and visions proliferate, but have to be aligned with existing organizational realities. After the montage, the chapter proceeds with a more formal introduction to the thesis, including its purpose and overarching research question. The chapter ends with an outline of the structure of the dissertation.

1.1 Prologue: following the data-driven organization

The vision of a data-driven organization – Flicking through PowerPoints

2017 – IT University of Copenhagen

I am in my office at the IT University of Copenhagen, located on the third floor of a modern glass building. In the atrium, several meeting rooms hover in the air, connected to the different floors of the building. The IT University itself is just as futuristic as many of the information technologies studied there. I am flicking through a PowerPoint presentation titled: *Data Driven Tax Administration*[1]. The author of the presentation is the Senior Vice President in the business area of Technology, Data and Security. According to the presentation, the DCTA, the backbone of the Danish welfare state, is pursuing a new vision: To become a data-driven organization. The presentation, which is shared online by The Senior Vice President, is published in 2016, and besides a few articles in the popular press and the reoccurring job applications that seek employees who can enable the tax administration to become data-driven, this is one of the few publicly available documents elaborating on the vision. Most pages in the presentation contain a headline, some very short sentences as bullet points, and a couple of images. One page in particular catches my attention. The headline states: data-driven detection, followed by seven bullet points. One of them is: "Customs → Machine learning to score risk on shipments to Denmark". Finally, on the bottom of the page it states that investments in "new tech and projects" cost 20-40 million Danish Kroner, with a potential revenue of 2-3 billion Danish Kroner - a return on investment of 50-150 times.

The risk score project – Meeting with the head of data analysis

February 2017 – DCTA: Headquarter

"I don't really like this notion of a data-driven organization. If you ask me, I would rather speak of an algorithmic-driven organization." The manager gives me a firm look before he continues: "First of all, I think there is a need to identify what it even means to be data-driven." This morning I avoided the 25-minute bike ride to the

[1] See (Overgaard, 2016).

university, and rode 10 minutes in the opposite direction, arriving in one of the headquarters of the DCTA. I am sitting in a meeting room together with the head of data analysis. I reply that I am interested in discussing what it means exactly, when the tax administration envisions itself as a data-driven organization. I ask him if it is possible to study some specific work related to this vision. He tells me about the risk score project, a project which aims to develop a machine learning algorithm to assist customs officers in selecting which imported packages should be inspected for tax fraud. I remember reading about this project in the PowerPoint presentation. The manager tells me that I can follow this project: "We do not need any help with writing code, we are already experts in this area. But I am interested in our relationship to the other parts of the organization." The manager tells me that he will set up the contact between me and the main data scientist working on the project.

The machine learning algorithm – Meeting with a data scientist

May 2017 – DCTA: Office for advanced data analysis

The data scientist points his finger at the computer screen – this is the machine learning algorithm that is going to enable the customs department and its employees to more efficiently and accurately decide which packages to inspect for tax fraud. I look at the screen, and see several lines of code consisting of text and numbers. The data scientist sits at his desk in the middle of an open office space. One row of desks after the other are lined up across the room. Headphones lie on the desks while some are placed on the ears of other data scientists working with their computers. The data scientist shifts to another tab on his computer and tells me that he is still working on improving the algorithm, so it is not yet fully developed. Although I have now 'seen' the machine learning algorithm, I am not entirely sure what it means.

The explanation algorithm – Participating in a workshop

May 2017 – DCTA: Customs department

In another branch of the tax administration, in the offices of the customs department, ten employees eagerly discuss what needs to happen in the next stage of the risk score project. Although everyone present is employed in the DCTA, in this context the employees from the customs department are considered customers with wants and

needs, and the three data scientists' task is to accommodate these needs. A project manager from yet another part of the tax administration facilitates the workshop. His job is to make sure that the conversation progresses and that everyone is heard. Late in the meeting, one employee from the customs department raises his voice: "I just thought about an additional need. It would be beneficial if the customs officers would get an explanation of why the algorithm selects a specific package for inspection." After a moment of silence, one of the data scientists replies: "The problem with some of these algorithms is that they are a bit black-boxed. That means that they are difficult, if not impossible, to explain. But we are in the midst of implementing another algorithm called LIME, which can provide an explanation of these machine learning algorithms."

The data bank – Reflecting on the workshop

June 2017 – DCTA: Headquarter

I am back at the headquarter of the DCTA. It's a stunning day and the sun is heating up the building, especially in those places where the façade is made of glass. I am meeting one of the data scientists. Before we head outside to find a place in the sun, we grab a cup of coffee. We sit down in a small yard, located in the middle of the building. The data scientist tells me that he was not that excited about the workshop. The project has run into some issues. The data scientist emphasizes that the data, which they had expected would be ready at this time, is still not 'flowing smoothly' between the different departments. A precondition for the project was that the data scientists would be able to receive data in real time from one of the new data infrastructures in the tax administration - the data bank. However, as the data scientist tells me, the data bank is still not ready to deliver data, which means that he will not be able to deliver any machine learning algorithm to the customs department. With those words, the data scientist looks at me and says with a defeated voice that he doesn't know if the risk score project is going to proceed for much longer, and that if he was to give me some advice, it would be to find another project for my research.

The machine learning algorithm – Receiving e-mails

February 2018 – DCTA: Customs department

This time it is an employee from the customs department who points his finger at a computer monitor. The employee is a former customs officer who is now employed to update and manage the rules in an IT-system that decides which imported goods and packages customs officers should inspect for fraud. Although the risk score project was closing down, the data scientist had still provided a machine learning algorithm, which the customs officers could test. The employee from the customs department tells me that he receives an e-mail from the data scientist every morning. He has just opened an attachment from the most recent e-mail, and he shows me a table with text and numbers – the output of the machine learning algorithm. The left side of the table displays an overview of packages about to be imported to Denmark. There is a score at the right side of the window that suggests whether or not the customs officers should inspect a package for fraud – the risk score. I ask the employee from the customs department whether they have had the chance to test the algorithm yet, and he replies that, so far, they haven't had the time to do so.

The inspection – Visiting the customs officers

July 2018 – DCTA: Copenhagen airport

I just arrived at the airport and find myself standing around a truck with several customs officers. The truck is selected for inspection. There is a vivid tension and excitement in the air as the custom officers start to pull down the fabric at the side of the truck, and several tree logs become visible. The previous day, another customs department had found a large number of cigarette packages hidden inside tree logs of a similar truck. It was because of that discovery that the truck today was selected for inspection. Even though we are standing several people around the truck, the possibilities for a thorough inspection are limited. A customs officer scans the outermost tree with a scanning device that looks like a flat baseball bat. Since the tree logs cover the entrance, it is not possible for the customs officers to look thoroughly at the logs furthest in the back of the truck. While one of the customs officers scans the tree logs, another is shining with a flashlight to look further inside the truck. With no wrong sounds from the scanning device and nothing to see with the flashlight, the

inspection is finished within 10 minutes, and the driver of the truck is allowed to continue.

After inspection – Talking to a customs officer

July 2018 – DCTA: Copenhagen airport

After the inspection, I talk to one of the customs officers. She is discouraged about the challenges customs officers face in actually performing inspections of imported goods and packages: "You saw for yourself what kind of facilities we have at our disposal, basically nothing. We can't climb onto the truck because if we get hurt, it's our responsibility. We can make an appointment and transport the truck to a storage facility in Copenhagen and inspect it out there, but that costs a lot of money and resources. Two employees have to drive from here and it quickly takes a whole day… We really wanted to rent the warehouse located just next to us. It's a big warehouse with ramps. Then we could empty the trucks here and now - that would be great. It's so unrewarding to walk around with the dogs out here as the trucks are too large for the dogs to help." Visiting the customs officers at the airport, I hear about the constraints and issues of the customs officers, who have been dealing with a large downsizing, and at the same time growing demands imposed on them by the European Union. After visiting and listening to the customs officers, I cannot help but to think about the 'distances' between different worlds of work. About the differences between the optimistic rhetoric of increased efficiency and high-tech solutions to the work of doing a customs inspection in practice.

Back at the IT University

July 2018 - IT University of Copenhagen

I receive an e-mail from the customs employee. He writes that the customs officers have now inspected some of the packages receiving a high risk score. In most cases, they didn't find any tax fraud. He also writes that they do not receive e-mails with the risk score anymore, and as far as he knows, the project has been transferred to another "IT-thingy" department. This is the last thing I hear about the risk score project and the development of the machine learning algorithm. I have become involved in

following another project in the tax administration, whose purpose is to develop a data-sharing infrastructure for sharing tax data with other public organizations.

1.2 Introduction & research question

The vision to become data-driven currently shapes the imagination and many practices in public and private organizations. Researchers proclaim that we have entered 'the era of big data' (boyd & Crawford, 2012) and that a data revolution is taking place (Mayer-Schönberger & Cukier, 2013; Kitchin, 2014). International organizations and Governments publish an endless number of strategies and reports discussing and promoting ideas such as big data, artificial intelligence, and the data-driven society (Danish Government, 2019; European Commission, 2018, 2020). In popular business magazines such as Harvard Business Review, one can read how those organizations that characterize themselves as data-driven outperform their competitors in terms of financial and operational measures (McAfee & Brynjolfsson, 2012). Central for many of these reports and statements are the technology companies Facebook, Apple, Amazon, Microsoft, and Google (Alphabet), who in Government reports and especially business magazines are promoted as role models for the 21st-century organization.

While the vision of the data-driven organization originated in the high-tech private sector, by now it has also become widely popular in more 'traditional' organizations. In Denmark, several public organizations, such as municipalities and health care organizations, re-imagine themselves as becoming data-driven (Høyer, 2019; Hockenhull, 2020). Although the data-driven vision is widespread in public organizations, recent ethnographic studies have highlighted that the vision simultaneously appears elusive and technocratic (Plesner & Justesen, 2020; Reutter & Spilker, 2019; Hockenhull & Cohn, 2021). Such studies have reported on managers and data strategists who struggle to define and adapt the data-driven vision into their specific organization, and data scientists and engineers who face legal and technical challenges as they experiment with new data technologies such as machine learning algorithms in public organizations (Reutter & Spilker, 2019). These studies question to which extent the emerging data technologies constitute a new way of governing and running public organizations or simply repeat existing attempts to quantify and automate work in an attempt to reduce government spending. This dissertation

contributes to the social inquiry of the data-driven organization with new empirical cases. It follows multiple calls to study ethnographically and conceptualize how digitalization and the current data-moment impacts work in public organizations (Kitchin, 2013; Plesner et al. 2018; Winthereik 2018; Møller et al. 2020).

As evident from the opening montage of this dissertation, the Danish Customs and Tax Administration (DCTA) has embarked on a journey to become a data-driven organization. However, inside the DCTA, the vision is also an object of debate and confusion. As the head of data analysis suggested to me, "First of all there is a need to identify what it even means to be data-driven" (See the prologue). This dissertation investigates the vision of a data-driven organization as both elusive and influential. It does so by attending to three different work situations in the DCTA. The thesis does not aim to settle the question of what it means to be a data-driven organization, but instead it aims to participate in an on-going debate about how current investments in and experiments with data infrastructures change public organizations. It views the data-driven vision not just as a technocratic ideal but as related to wider transformations of work and organizational relations. As highlighted in the opening montage, the data-driven vision is often presented rather technocratically as the implementation of new data technologies and infrastructures such as machine learning algorithms, data platforms and distributed databases. Technologies imagined making public organizations more efficient and increasingly able to provide better services for its citizens. This dissertation will move beyond the technocratic discourse and direct attention to work carried out in a public organization working to become data-driven. It qualifies abstract notions, like big data and algorithms, by investigating "how public sector employees deal with the current digitization imperative in their work practices" (Plesner et al. 2018, p. 1187). The dissertation explores and answers the following overall research question:

Research Question

How are work situations reconfigured as the Danish Customs and Tax Administration strives to become data-driven?

Studying the reconfiguration of work situations in the DCTA means to study the material and semiotic effects of the organization's current investments in an attention to data. The notion of reconfiguration, which I take from Lucy Suchman, will be

further discussed in chapter 3. It points to the, "on-going elaboration and potential transformation of culturally and historically specific practices" (Suchman, 2007). I study the reconfiguration of the tax administration with special attention to shifting organizational relations and boundaries between employees within the tax administration, between tax employees and taxpayers, and between the tax administration and external actors such as IT suppliers.

I address the research question more specifically through analysis of three different work situations in the DCTA. In each of these work situations, public employees were working with the implementation of new technologies or technology-based visions. In the first work situation, I direct attention to frontline workers at the 'margins of data infrastructures', whose work has become somehow secondary to the self-service platforms, that ideally make it possible for taxpayers to manage their taxes without any help from tax personnel. In the second work situation, I analyse the development of a machine learning algorithm to help customs officers detect fraud in packages imported to Denmark. The machine learning algorithm was imagined to be more accurate than the current algorithms for selecting which packages customs officers should inspect for possible fraud. Finally, the third work situation I analyze is the development of a data-sharing infrastructure imagined to foster a cheaper sharing of tax data between public organizations, and to provide tax employees with more responsibility for data sharing.

In each of these work situations, I found a particular figure which was central to the work situation and for the public employees working with a particular digital transformation. In the first work situation, it was the figure of the 'no-touch customer', which had become the preferred taxpayer ideal. This figure did not just designate an ideal image of taxpayers, it also affected the interaction between frontline workers and taxpayers. In the second work situation, I found the figure of 'the customs officer's nose'. This figure was central when data scientists and customs employees contemplated about how data infrastructures aligned with the customs officers work. The central figure in the third work situation was that of 'bringing home'. The project participants used this figure to think about relying on employees from the tax administration as much as possible in developing a new data infrastructure.

For me, these figures are central for understanding the effects of new data-driven visions and technologies. They interrelate and interfere with organizational issues and the implementation of new technologies. They open up the posing of analytical as well as political and organizational questions, such as how boundaries and responsibilities get reconfigured in the tax administration. The figures trouble the perceived certainties of the situation (Haraway, 2018). Like Donna Haraway's conceptualization of string figures (Haraway, 2016), they show how transformation processes are weaved into historical and material conditions. In this view, working with digital transformation requires taking the existing organizational patterns, which are handed over from earlier work, seriously in order to understand what new patterns may possibly be weaved.

1.3 Structure of dissertation

The dissertation consists of two parts. The first part provides an overall framing of the dissertation, including a conclusion which summarizes the three research papers and discusses some cross-cutting themes. The three research papers are presented in their full lengths in the second part of the dissertation. The first part of the dissertation proceeds as follows:

Chapter 2 gives a brief organizational background to the Danish Customs and Tax Administration (DCTA). It highlights overarching organizational changes, which have taken place since the beginning of the 21st century, and shows how organizational restructurings and visions have been inextricably linked to the development and use of IT-systems. The chapter also highlights the emergence of the vision of a data-driven tax administration, which has entered the DCTA in the second decade of the 21st century.

Chapter 3 introduces the analytical perspective of the dissertation. Drawing on Science and Technology Studies, it highlights an analytical orientation towards studying the work involved in constructing a data-driven organization. The chapter describes the study of the reconfiguration of work situations, which is inspired by the work of Shoshana Zuboff, Susan Leigh Star, and Lucy Suchman.

Chapter 4 provides an overview of the fieldwork which was carried out between 2016-2018. It presents the methods used to collect and analyze the ethnographic material of the dissertation. Drawing on actor-network-theory and situational analysis, the chapter describes how fieldwork in the DCTA is approached. The chapter ends with reflections on the access difficulties I encountered during the fieldwork and what these access difficulties can tell us about the organization and the ethnographic study of data infrastructures.

Chapter 5 provides a brief overview of recent academic literature concerned with the social, political, and organizational implications of intensified investments in and attention to data. The chapter introduces studies of data-driven organizations, which among other things, highlight the invisible work supporting data-driven organizations. The chapter ends with a presentation of the research and perspectives on how the vision of data-driven organizations has impacted public organizations.

Chapter 6 discusses the vision of the data-driven organization and highlights the findings of the dissertation. It highlights the arguments and findings in the three research papers, and subsequently it discusses some general analytic points and themes across the three research papers.

2 ORGANIZATIONAL BACKGROUND

This second chapter provides a brief organizational background of the DCTA. The aim of the chapter is to highlight how organizational reforms and changes in the DCTA have been inextricably linked to visions of new technologies and accompanying techniques. It situates the emergence of the data-driven vision within a longer trend of digitalizing the tax administration. The data-driven vision and the latest technologies it promotes are often presented in a somewhat ahistorical manner, focusing on the future rather than the past. However, as this contextualizing section will show, the development of IT systems does not appear in a historical vacuum.

Figure 1. Queue at the tax office in Odense in 1963 (Source: Fyens Stiftstidendes pressefotosamling, Odense Stadsarkiv)

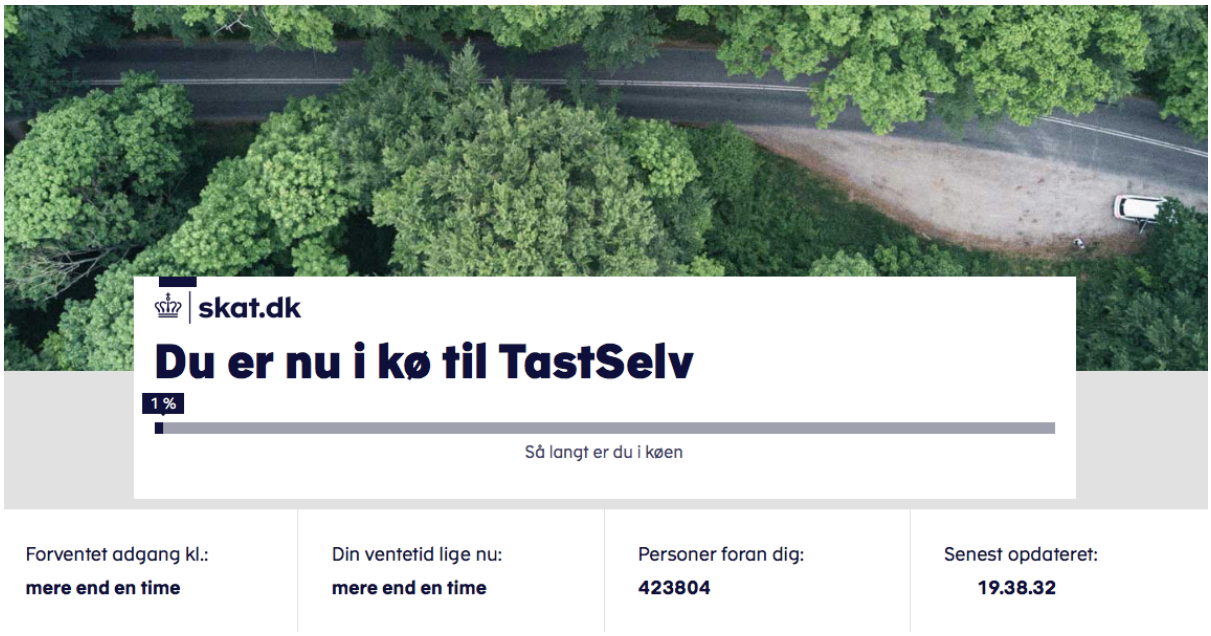


Figure 2. 'You are now in the queue for TastSelv' (own photo 2021)

2.1 The Danish champion of digitalization

It is self-evident that a well-functioning tax administration is essential for the Danish welfare state. The DCTA plays a central role in Danish society collecting the equivalent of 134 billion euros in yearly taxes, which amounts to 50% of the Danish gross domestic product. The income from tax supports Denmark's central infrastructures, such as the education system, the health care system, the police, the defense, and a comprehensive welfare system. According to several studies and reports, the Danish taxpayers generally support paying high taxes, as long as they are used to finance the central welfare institutions (Jessen, 2018; Keiding, 2019).

Since the 2000's the DCTA and the administration of taxes have changed in significant ways. One illustration of these changes is the image of the queue (see figure 1 & 2). While it has become the norm that Danish citizens administer their taxes online, they still find themselves waiting in queues at their telephone or their computer. This shows how, in general, communication with the tax administration has changed since the turn of the millennium. Instead of waiting in a physical queue at a municipal building, taxpayers today wait in queues on the telephone and on their computers, to access systems or talk to administrators. Taxpayers who access their annual tax return in the immediate days after its release can find themselves in a digital queue for 'more than an hour' alongside more than 400.000 other Danish taxpayers (see figure 2). The online queue reminds us that even digital infrastructures have their own constraints and limitations.

From the Danish taxpayers' perspective, receiving and delivering tax documents online is probably the most obvious and visible way digitalization has affected the DCTA. When Danish taxpayers view and edit their tax documents online, they use the IT-system named TastSelv[2]. TastSelv was first implemented in 1994, where it became possible for taxpayers to report their taxes through the telephone, and in 1995 this service was extended to the internet. Today, taxpayers by default use this online system to manage their taxes, unless they have received a formal exemption from the Danish state – a policy that follows the overall Danish digitalization strategy that citizens are 'digital by default' (Schou & Pors, 2018). For taxpayers, TastSelv can

[2] Directly translated to English TastSelv means 'type it yourself'. In English the DCTA refer to the system as e-tax, but throughout this dissertation I use the Danish designation of TastSelv.

appear relatively simple as they interact with a well-designed user interface. But the historical developments which eventually made it possible to develop the TastSelv system are far from simple. The history, which eventually made it possible to implement the TastSelv solution, includes politicians with grand visions, the establishment of new IT organizations, investment in technologies, and partnerships between public and private organizations (Heide, 1994; Danish Ministry of Taxation, 2004; Johansen, 2007; Østergaard, 2015).

In the 1950s, the Social Democratic Minister of Finance, Viggo Kampmann, pushed for more electronic data processing in the tax administration. He advocated for implementing a pay-as-you-earn system (PAYE), which he viewed as essential for ensuring the future taxes for the welfare state. PAYE is the idea that taxes are deducted by the employers before the salaries are paid to employees. In conversations with the IT company IBM, Viggo Kampmann found that two elements were central for creating an effective PAYE system: A central register for all taxpayers and a central register for employers (Østergaard, 2015). In 1959, the Danish Government established I/S Datacentralen, a centralized IT organization for developing and operating administrative IT systems for the Danish state and municipalities. The managing director of Datacentralen became Willy Olsen, who had been employed at IBM as the director of sales to the Danish Government, and who had played a central role in consulting Viggo Kampmann about how to develop the PAYE system (Prosa, 2014). Datacentralen installed its first IBM mainframe computers in 1962, and in 1964 the central register for employers was established, in 1968 the central register for taxpayers was established, and in 1970 the PAYE system was developed (Prosa, 2014). The two registers, which today are known as the CPR and CVR registers, are often highlighted when explaining the success of Denmark's digitalization.

Much of the work involved in the development of the PAYE system paved the way for establishing the TastSelv solution. This includes the development of the central registers for employers and citizens, and the data-sharing agreements between the tax administration and employers, banks, and pension funds. The historical events which led up to the development of TastSelv illustrate some of the complexities that can surround the development and maintenance of IT systems. From a user perspective, IT systems can seem quite simple, as if they have appeared overnight, detached from

any historical, political, and organizational context. However, the preconditions for establishing TastSelv included setting up national registers, establishing public organizations, and developing new routines for reporting on taxes. Developing IT systems is not just about introducing the right hardware and software, it is about shaping the environment in which IT systems are imagined to function.

In the years following the implementation of TastSelv, the DCTA was praised nationally and internationally. With TastSelv, taxpayers could not only access their taxes online, but a large part of their tax information was also pre-filled by the tax administration in collaboration with third party organizations (OECD, 2013, p. 241). The DCTA won the Danish e-business price in 2004 for TastSelv (Østergaard, 2015), and in 2006 it was awarded the prize as the Danish champion of digitalization (Bang, 2006a). The DCTA received the award as the champion of digitalization due to its ability to handle "a large number of digital transactions", and due to how the organization had improved its services to taxpayers through information technologies (Bang, 2006a). After being awarded as the Danish champion of digitalization, a strategy director explained that the tax administration was yet to reap all the benefits from digitalization. He foretold that by 2010 the DCTA would have halved its number of employees due to their investments in digitalization (Bang, 2006b).

The strategy director's promise of halving the number of employees in the DCTA was part of a more extensive reorganization of the tax administration, which had been announced in 2004 by the Ministry of Taxation. An organizational centralization would merge the central tax administration with the municipal tax centers into a unified tax administration under the responsibility of the state. As a consequence of this reform, 276 municipal tax centers would be closed. The reform was based on the idea that an organizational centralization and the development of new IT systems would provide large efficiency gains and allow the tax administration to drastically lower its number of employees. Therefore, it was no secret from the beginning that the plan depended on the successful development of a range of new IT systems.

> To establish the new Ministry of Taxation is a large and comprehensive task. The changes will take some time. Many of the previously mentioned elements presuppose new and improved IT systems. These will be acquired and modernized in the next 1 to 5 years. (Danish Ministry of Taxation, 2004, p. 10)

2.2 Crisis

The 21st century began with optimism in the DCTA. The IT-system TastSelv had been a huge success, and the development of a centralized tax administration promised a more efficient administration of taxes. However, the optimism and the high expectations for the unified tax administration faded during the following decade, and instead the time period between 2005-2015 became known as one of rashness and powerlessness (Christensen & Mortensen, 2018). The tax administration did reduce its number of employees from 11.500 to 6.500 by 2014, but the reduction of employees came with negative implications and consequences.

First of all, several IT systems were not realized as planned. In 2007, a large setback for the digital champion occurred as the implementation of an IT system, which was fundamental to running the new digital tax administration, was postponed to 2009. The IT-system was called EFI, and the purpose of the system was to automate the collection of public debt. The postponement of EFI was problematic for several reasons, not least because many of the employees who had been responsible for collecting debt in the municipalities had been let go or relocated to other positions in the tax administration (Borre et al., 2017). It was no longer easy to return to the manual collection of public debt. The DCTA had thus implemented organizational changes before the promised benefits of the new IT systems had actualized. As EFI was postponed, the public debts grew. In 2009, EFI was still not functioning properly and it was postponed for another four years. In 2013, the Minister of Taxation at the time from the Socialist People's Party, decided to put EFI into use although several had warned that the IT-system still had issues. In 2015, the Legal Adviser to the Danish Government reported that the system didn't function properly. It was collecting a large amount of expired debt and was thus operating illegally. The message from the Legal Adviser was clear, EFI had to be terminated (Danish Ministry of Taxation, 2015a).

EFI is often highlighted as a prime example of the crisis of the DCTA, but it was only one out of several IT projects that followed the pattern that Bent Flyvbjerg has called the 'iron law of megaprojects': "Over budget, over time, under benefits, over and over again." (Flyvbjerg, 2017). Six IT projects part of the 'system modernization' of the tax

administration, including EFI, ran over time and over budget. The system modernization was, as already noted, a central part of the plan to reduce the employees of the tax administration by 40% (Christensen & Mortensen, 2018, p.135). When the National Audit Office of Denmark in 2015 evaluated the system modernization, they reported that it had in total become three times as expensive as planned, and was delayed by 5 years (Christensen & Mortensen, 2018, p. 146). The National Audit Office of Denmark made a thorough critique of the DCTA's economic management of the projects, and of its high willingness to take risks. The Minster of Taxation at the time responded that the critique should also be levelled at the IT-suppliers who had been hired by the DCTA (Christensen & Mortensen, 2018).

The difficulties associated with the development of new IT-systems were not the only challenges the tax administration faced in the period between 2005-2015. Many cases of tax fraud received attention in the Danish media (Aagaard, 2017). The most serious of these cases was fraud in relation to the reimbursement of dividend taxes (In Danish: *udbytteskandalen)*, which has been called the largest scam in the history of Denmark (Holst, 2016). Denmark was one out of several European countries that had been scammed by a group of international fraudsters. In Denmark, the fraudsters managed to trick the tax administration for almost 2 billion euros between 2012-2015, by applying for reimbursement taxes based on fabricated documents (Danish Ministry of Taxation, 2020). They had managed to infiltrate the tax administration, and received help from one employee who had become solely responsible for approving reimbursement applications to foreign companies due to the downsizing of the tax administration. In a commission of inquiry which followed the scandal, the manager with responsibility for the reimbursement area explained that the tax administration simply didn't have sufficient resources available to do thorough inspections of the companies requesting reimbursements (Elkjær, 2019).

The many scandals did not just lead to a large amount of tax money being lost, it also damaged the reputation of the tax administration and led to a declining trust in the organization from Danish taxpayers. In 2014, the Ministry of Taxation reported on this declining trust, based on its customer satisfaction surveys (Danish Ministry of Taxation, 2014). In 2015, the newly appointed Minister of Taxation from the Liberal Party of Denmark called for political unity and for collaboration across the political

spectrum. Between 2005 and 2015, the DCTA had nine different ministers of taxation from four different political parties, which meant that the responsibility for past mishaps could not be ascribed to a single government or political party. The new Minister of Taxation encouraged everyone, including the Danish media, to focus on the future and to work together to get the tax administration out of its crisis (Danish Ministry of Taxation, 2015b). In 2016, the Minister of Taxation announced a new large reorganization of the tax administration, highlighting the lack of trust as a central concern and thus as a central reason for the reorganization (Danish Ministry of Taxation, 2016). In the report, *A new tax administration – new organization, more employees, and IT that works*, the Minister of Taxation announced massive investments in both increasing the number of employees and improving IT systems for the tax administration.

> It is important in this regard to make it clear that we have learned from our past mistakes. Our previous work with EFI should not make us afraid of digital solutions. Digitization is and will be a part of the future (Danish Customs and Tax Administration, 2016a, p. 2).

2.3 IT as problem and solution – Becoming a data-driven tax administration

Although the DCTA between 2005-2015 experienced several scandals, directly and indirectly related to failures and challenges with IT systems, the Minister of Taxation's proposal to restructure the tax administration made it clear that IT was an inevitable part of the tax administration's future – this time the Minister emphasized that the tax administration would invest in IT that works (Danish Ministry of Taxation, 2016a).

In the two reports from 2015 and 2016, the Ministry of Taxation highlighted several of the challenges associated with the tax administration's IT infrastructures.

> DCTA was earlier a leader in the development of IT in the public sector. Today DCTA has a series of basic challenges with its use of IT. DCTA has around 200 systems, which have been developed ad hoc and are often highly interdependent. Central IT-systems are supplied by different external suppliers. Knowledge about data and business processes, which are 'moulded' into the IT systems, are to a large extent placed with the DCTA's IT-suppliers and not with the DCTA itself… These structural challenges inhibit DCTA's possibilities for managing and controlling the development of its use of IT. (Danish Ministry of Taxation, 2015b, p. 5)

One of the key challenges identified was a complex IT landscape, which had left employees in the tax administration with little knowledge and control over the IT systems, and reflected a general lack of IT competencies within the organization. The tax administration problematizes its 'digital heritage' and its 'legacy-systems', referring to the fact that many of the central IT systems of the tax administration are based on IT from a different time stretching back to the 1960's (Danish Ministry of Taxation, 2015b, p. 25). Another main challenge addressed in the reports was the difficulties of managing and running IT projects. The difficulties are associated with IT systems that are too large and driven by external suppliers and consultants.

As a response to these challenges, the Ministry of Taxation described some clear changes, which delineated the new approaches from those of the past. Rather than investing in large IT systems, the tax administration would strive to divide IT systems into smaller parts to be developed individually (Danish Ministry of Taxation, 2015b, p. 25). Rather than relying on traditional waterfall development methods, the tax administration would begin to use more iterative and agile software development methods. Rather than relying heavily on outsourcing IT solutions to private companies, the tax administration would take on increased responsibility for the management and development of IT systems, and in general improve their internal IT capacities.

This new approach to IT did not just involve rhetorical shifts. In 2016, the DCTA established a new business area called IT & Data, described as a first step in creating 'a larger managerial focus' on IT and data (Danish Ministry of Taxation, 2016b). Consequently, in 2017, the Minister of Taxation announced that by 2018, seven independent agencies would replace the unified tax administration that had been established with the structural reform in 2005 (Danish Ministry of Taxation, 2017). In this new major organizational reform, the IT and development agency (*Danish: Udviklings- og forenklingsstyrelsen*) was established as the third largest agency. The agency employs around 1500 employees distributed across five departments and approximately 80 separate offices (see figure 3). The establishment of the IT and development agency is a highly visible manifestation of the DCTA's attempt to take increased responsibility for its IT solutions.

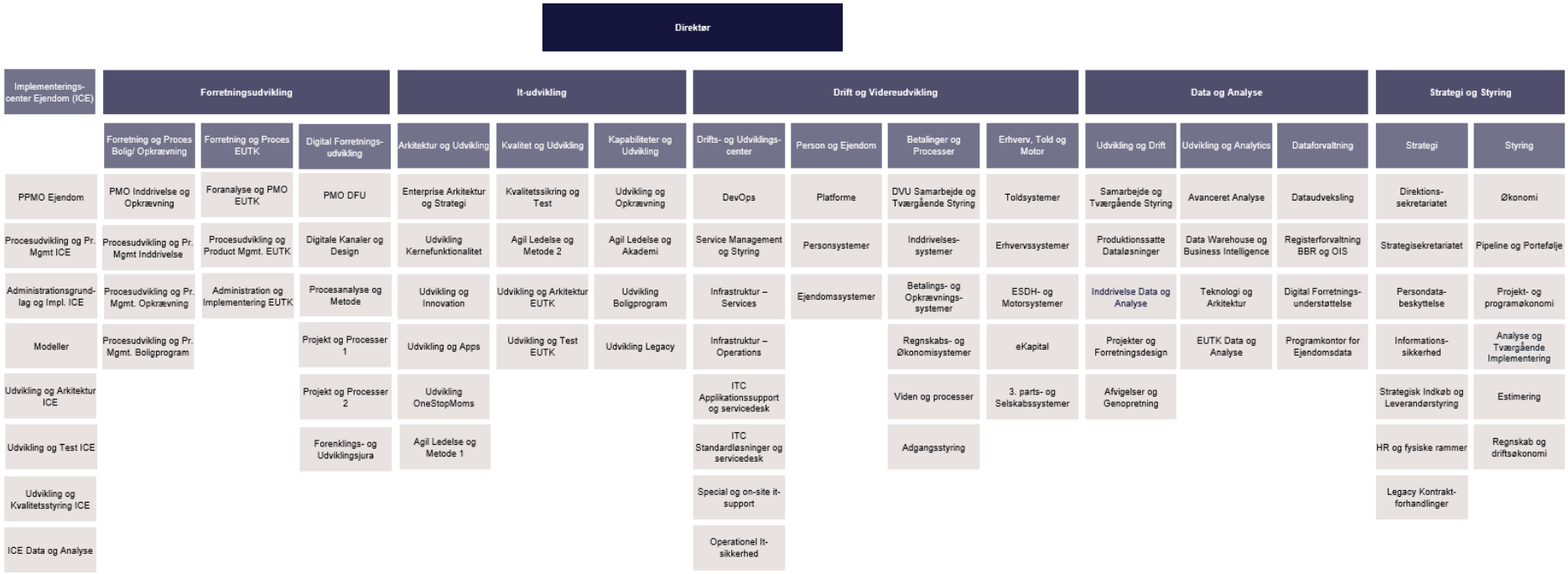


Figure 3. Organizational chart of the IT and Development Agency (Source: https://www.ufst.dk/om-styrelsen/organisation/)

At the same time as the DCTA restructured its organization, and IT gained a more prominent organizational role and visibility, a new vision was emerging in the tax administration – to become a data-driven organization. In 2016, when the Ministry of Taxation announced that IT & Data would be established as its own separate area of business, the manager of the new business area announced that, "with a unified IT & data we will sharpen our focus on the important task of creating a new data-driven tax administration, where the use of relevant data increasingly will be a part of understanding our customers needs" (Ministry of Taxation, 2016b).

Earlier in 2016, the idea of the data-driven tax administration had been promoted in the Forum on Tax Administrations (FTA), an OECD forum that brings together commissioners from 53 different countries. The FTA published three reports in 2016, in the *tax administration of the future series* (OECD, 2016a; OECD, 2016b; OECD, 2016c), discussing how tax administrations actually and ideally should use digital technologies in their operations.

> Bringing a data-driven culture into a tax administration implies developing a culture in which data and analysis drive all aspects of the organisation. This is a journey not a destination. It will continue as new technologies emerge; digital disruption creates even greater innovation and customer expectations change (OECD, 2016b, p. 58).

The reports highlighted technological innovations such as cloud computing, blockchain, big data, and the internet of things, as they explored the key themes of changing customer expectations and the role of emerging technologies.

With new organizational visions, and a large re-organization including investments in IT competencies and technologies, the Danish Ministry of Taxation has responded to the previous decades of IT failures and challenges. While this dissertation does not attempt to predict the success or failure of the new data-driven tax administration, it provides insights into some of the tentative work taking place in the DCTA, and highlights some of the organizational tensions, which are likely to pose a continuous challenge for the tax administration throughout the coming decades.

3 THEORETICAL ORIENTATION

This third chapter presents the dissertation's theoretical underpinnings. First, the chapter outlines the overall theoretical orientation towards work done in the DCTA to become a data-driven organization. This orientation towards work is guided by sensibilities taken from Science and Technology Studies. Second, it highlights the work of Shoshana Zuboff, Susan Leigh Star, and Lucy Suchman as main sources of inspiration. With inspiration from these scholars, I draw attention to what I call the reconfiguration of work situations, which become visible upon studying public employees working with data infrastructures.

3.1 Visiting the construction site

The cover of, *Reassembling the social – An introduction to actor-network-theory,* pictures a construction site (Latour, 2005). In the foreground, two men are discussing with one another while pointing at the site. In their hands, they hold the plans and visions for what is in the midst of being constructed. In the background, we can see the contours of two unfinished buildings. Workers are climbing up and down ladders and setting up the wooden foundation of the buildings. To begin this chapter, I use this construction site image to draw out some sensibilities from STS, which have guided my study of the data-driven tax administration as a site that is also under construction.

First, the construction site image illustrates an ambition to study the mundane work required to build and maintain infrastructures such as buildings, organizations, information technologies, and even science. Several key texts in STS have made work a central theme of analysis. In what are known as laboratory studies, several researchers 'opened the black box' of scientific knowledge and analyzed in great detail how facts were constructed through meticulous and laborious work (Cetina, 1995; Latour, 1987). As pointed out by Susan Leigh Star, "a central tenet of the pragmatist work in STS has been to think of scientists as people who are doing a certain kind of work" (Star, 1995a, p. 15). The laboratory studies aimed to provide an alternative to idealized descriptions of science, as adhering to clear methodological principles , or reliant on the extraordinary cognitive capacities of individual geniuses (Blok & Jensen, 2009, p. 52). Instead, laboratory studies highlighted the organizational and material conditions, involved in the construction of scientific knowledge. This shift has been described as a shift in attention from epistemology, a focus on what counts as true knowledge, to ontology; the question is not how does science describe the world, but what worlds scientific practices participate in bringing forth (Mol, 2002). Rather than viewing entities such as science, organizations, and information systems as pre-existing entities with fixed properties, STS studies direct attention to the continuous work required to uphold these entities. STS scholars thus draw attention to the process of organizing rather than the formal properties of organization (Law, 1994; Czarniawska, 2004; Latour, 2013).

Second, the image illustrates an ambition to attend to and describe a wide range of entities and materials involved in a work situation. On the construction site image, we

see human workers, plans and visions for the construction site, and the materials involved in the construction process – such as wood and cement. One of the often-highlighted ANT hallmarks is the point that both human and non-human actors participate in construction work (Pickering, 1992). In some studies, this attention to non-human actors manifests itself through technologies that 'speak' (Latour, 1996), non-humans that act (Callon, 1984), or more simply an interest in technical and material entities as a central part of social studies (Star, 1999). The main point is that societal changes are not just a matter of human intentions and social processes, but that a wide variety of non-human actors shape this process. 'Agency' is diversified and pluralized. STS researchers have used different concepts to designate and speak about the mix and co-constitution of human and non-human actors, such as actor-networks (Latour, 2005), heterogeneous engineering (Law, 2012), socio-material assemblages (Orlikowski & Scott, 2008), and human-machine configurations (Suchman, 2007).

Finally, the construction site image illustrates how representations of work depend on a particular staging of the relationship between foreground and background. As observers of the image, we clearly see the two managers at the front holding the plans and visions in their hands. Almost out of sight are the workers in the background who climb around on the wooden structures. STS studies are attentive to the different ways in which what counts as foreground and what count as background in a situation may be staged. Some studies in STS follow a normative inclination of bringing the work that is left out of popular discourse into view, highlighting the backstage work required to uphold a foreground (Star, 1991; Law, 2004).

In the following section, I present three theorists who have been central inspirations throughout the work on my dissertation . These researchers have all explicitly studied the interplay between organization, work, and information technologies through ethnographic methods.

3.2 Work, IT, and organization – Three guiding perspectives

Shoshana Zuboff, Susan Leigh Star, and Lucy Suchman have particularly shaped and guided my study of how a public organization and the work of public employees are affected by data-driven technologies and related visions. All three conducted workplace ethnographies in the 1980s and 1990s, where they described and analyzed

how computers, and later the internet, impacted organizations and workers. They view the relationship between work, technology, and organization as being mutually constituted. Whereas I am interested in the vision of the data-driven society, their studies from the 1980s and 1990s worked with and against the notions of the information society and knowledge society (Drucker, 1969; Bell, 1974). In the following, I present some of the key ideas and concepts of these authors, which I draw upon and discuss in this dissertation.

3.2.1 Shoshana Zuboff – Information technology and the changing nature of work

In her book from 1988, *In the age of the smart machine – The future of work and power* (Zuboff, 1988), Shoshana Zuboff analyze how the spread of computers affected American workplaces in the 1980's. In the preface to the book, she writes about her research intentions in the following way:

> I wanted to discover the flesh and blood behind the concepts, the interior texture rather than the external form. I wanted to understand the practical problems that would have to be confronted in order to manage the new computerized workplace in ways that would fulfil the lofty promise of a knowledge-based society. (Zuboff, 1988, p. xiv)

This dissertation shares this research intent directed at the similarly lofty promises of the data-driven society. Although Zuboff carried out her research in the 1980s, there are striking resemblances between the discourses and empirical details she presents in her ethnography, and what can currently be witnessed as a data revolution. For example, she highlights how the adoption of new information technologies, such as algorithms and data-base systems, are imagined causing rapid and all-encompassing changes for both organizations and individuals.

Zuboff conducted ethnographic studies in seven American organizations exploring how IT was changing the workplaces, including the relationship between workers and the relationship between workers and management. Zuboff contends that there is no simple way to explain the relationship between IT and workplace transformations. How organizations are affected by IT depends on the industry, the historical conditions of the company, and the existing organizational structures and relations. She uses the metaphor of the Kaleidoscope to argue that small changes in how organizations use and apply IT, corresponds to turning a Kaleidoscope making new

and complex patterns visible (Zuboff, 1988, p. 387). While IT has the power to generate new organizational patterns, Zuboff stresses that it is the human hand that makes the kaleidoscope turn, arguing that human choices are essential for deciding which patterns to create.

While there is no simple answer to explain how IT affects workplaces, Zuboff suggests that IT systems provide a specific organizational opening – an opportunity to redefine and rethink organizations and their work relations. She argues that information technologies are unique in comparison to earlier technologies of the industrial age because of their capacity to informate. Zuboff writes that IT is both similar and different from previous technologies, in that it both has the capacity to automate and informate (Zuboff, 1985). On the one hand, she views IT as a continuation of technologies of the industrial age as it automates tasks previously carried out by workers. On the other hand, she views IT as something new because of its ability to informate. It offers information or data about the process that the technology controls or manages. In this sense, she writes about IT having a reflexive dimension: "Information technology not only produces action but also produces a voice that symbolically renders events, objects, and processes so that they become visible, knowable, and shareable in a new way" (Zuboff, 1988, p. 9).

Zuboff describes how the duality of IT, and especially its informating ability, have implications for the work practices of organizational members. Information technologies alter work in the manufacturing industry such as pulp and paper mills. Mill Operators experience how their work becomes displaced. Instead of working directly with the industrial machines, they become machine operators working at a computer terminal. In this new situation, operators have to be able to explicitly state how the pulp and paper machines operate in order to make adjustments to them through a computer interface in contrast to earlier where the operators knowledge about the machines were located in their bodies and senses. She describes how workers more or less stopped complaining about muscle and back pain and instead complained about getting headaches from working intensively at computers. She describes the introduction of computers as a transition from action-centred skills to intellective skills (Zuboff, 1988, p.62).

Besides industrial workers, Zuboff also addresses how managers and clerks were affected by the introduction of personal computers and data-base systems. She writes about a bank implementing a new data-base environment. The data-base was imagined storing the knowledge of managers and clerks, such as information about customers and work tasks. Information and knowledge which managers had sometimes written on paper, lying on their desks or in their drawers, would all become part of a database, potentially for every other organizational member to access. Zuboff argues here, that information technologies lead to a convergence of different forms of organizational work which become more alike. As most employees come to work with computers, there is no longer a clear distinction between what was once referred to as blue-collar workers and white-collar workers. As organizational knowledge becomes displaced from the individual minds of workers to centralized databases, knowledge becomes publicly available, rather than being private.

For Zuboff, this homogenization of work, creates a demand for rethinking existing hierarchical organization structures and to redefine authoritarian relationships in the organization. The central question for Zuboff is whether workers become empowered and organizations more democratic – or, to the contrary, if workers become disempowered and organizations more authoritarian. As noted by Kallinikos, "a key claim of the book pivots around the necessity of finding alternatives to centralization and hierarchy" (Kallinikos, 2010, p.1100). Although information technologies make workers more alike, Zuboff claims that the informating capacities of IT can be used to reinforce existing boundaries if the access to information is tightly regulated and centrally controlled. Zuboff ends with proposing a vision of the informated organization. The informated organization uses the informating potentials of IT to create a more democratic and egalitarian organization, which emphasizes the sharing of information and organizations as learning institutions.

While Zuboff concluded her work, *In the age of the smart machine,* with her positive image of how computers could contribute to the flourishing of organizations and their workers, she has in her recent work made it clear that today's organizations have mostly exploited IT's authoritarian capacities. In 2019 Shoshana Zuboff published the book *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Zuboff, 2019). This book, which has become highly popular, critiques the

emergence of the data-driven society and the tech companies that profit. In a dystopian analysis, Zuboff highlights how the prominent technology companies, such as Facebook and Google, reproduce and accelerate existing capitalist logics. Surveillance capitalism describes a logic where technology companies are incentivized to accumulate as much data on people's lives as possible. Technology companies do so by creating monopolies of information, and keeping users hooked on their platforms and services. Rather than open and publicly available databases, these technology companies rely on their proprietary access to data, and an asymmetrical relationship to their customers. The technology companies not only collect data to analyze their customers and users but to nudge and control them. She argues that where earlier forms of capitalist organizations involved a colonization of the public space, the new companies of surveillance capitalism colonize the most intimate spheres of private lives and our thoughts.

3.2.2 Susan Leigh Star – Information infrastructures' visible & invisible work

Susan Leigh Star was a sociologist with a background in the school of symbolic interactionism (Blumer, 1969; Becker, 1982). She describes herself as an STS scholar, and served as the President of the Society for Social Studies of Science between 2006-2007 (Balka, 2010). According to Star, a fundamental task of sociology is to attend to work (Star, 1991; Star & Strauss, 1999). Star approaches the study of work from a different perspective than Zuboff. She investigates what counts as work in specific empirical situations. She is interested in what work gets appreciated and recognized, and what work becomes relegated to the background, not captured by official discourses and articulations.

> Who are *all* the people working on a given production? How does it change our understanding of art, or science, or technology, when we restore all the work?... To do a sociology of the invisible means to take on the erasing process as the central human behavior of concern… This is in the end a profoundly political process, since so many modern forms of social control rely on the erasure or silencing of various workers, on deleting their work from representations of work. (Star, 1991, p. 281)

Rather than situating her study of work in relation to sociological concepts, Star often begins her studies reflecting on everyday understandings of work. One of Star's stronger definitions of work is as the glue which bridges invisible and visible work (Star, 1991). For example, she highlights that visible abstractions, such as

representations of a work situation or an academic theory, are the end product of a lot more work. But once abstractions are constructed and settled, it also requires work to challenge them and/or to reconnect them to the practices they purport to describe. The point is not that abstractions are inherently good or bad, but that they should be seen as a topic for sociological investigation.

> The oxymoron contained in the label Grounded theory is a clue that this method is a way of wrestling with that which joins the visible ground with the invisible abstraction. That the 'glue' is work appears in various ways throughout the discussion, focusing on the work of the analyst and the work done by the people being studied. (Star, 1991, p. 270)

Star argues for the importance of attending to the work involved in science, organizations, and the development of IT on practical, political, and moral grounds. On a practical level, she argues that it is fundamental to understand the complexities of work in order to make useful interventions and changes to work practices. Misunderstanding or underestimating work leads to bad decisions and outcomes, such as failures related to the implementation of organizational strategies and the development of IT systems. Information systems are often introduced with the promise of automating or supporting parts of existing work. If work is mischaracterized or underestimated in the attempt to develop and implement IT systems, they will be likely to fail[3].

Star also attends to work as a moral and political issue. She suggests analyzing work with particular attention to marginalized positions. As Kjeld Schmidt writes about her:

> One of the things that characterized Leigh's life and work was a determined effort to have the skills and concerns of ordinary workers respected and taken into account in the design of technical systems, in organizational management, and in public discourse. (Schmidt, 2016, p. 345)

She highlights that although formalized systems, such as information systems, might be efficient for a large group of people, they also tend to exclude certain groups of

[3] Drawing on Star's work, Diana Forsythe argued that the implementation of artificial intelligence, which she studied, failed because the computer scientists weren't able to incorporate the tacit knowledge of experts whose work should be replaced. Forsythe argued that AI-based IT systems, for this reason, were likely to be decontextualised (Forsythe, 2001).

people. One of her popular examples comes from a personal anecdote about being allergic to onions (Star, 1990). As a consequence of this onion allergy, she experienced waiting 45 minutes on a burger from McDonalds, because she ordered a burger without onions. Star uses this experience to describe that highly efficient systems can be inefficient in dealing with certain edge-cases. Indeed, very efficient and formalized systems are prone to produce marginalization. Highlighting these inefficiencies can be a way of developing more inclusive systems, and to notice and discuss the shortcomings and political consequences of existing ones.

Star investigates work in a wide range of settings ethnographically. Like many STS scholars, she was originally interested in the work of the scientist. She was particularly interested in how scientists collaborated and communicated with each other, as well as with groups with other types of expertise (Star, 1985; Star & Griesemer, 1989). Star argues that successful collaboration does not depend on rational and clearly defined terms of communication. Instead, this relies on what she termed 'boundary objects'. According to Star, boundary objects are central for interaction across different social worlds or groups, while allowing the different social groups to interpret the object differently and relate to it differently in work practices. Star develops the concept of the boundary object in an article from 1989 to describe how different professions at the Museum of Vertebrate Zoology could cooperate while maintaining their different viewpoints and expertise (Star & Griesemer, 1989).

In the 1990's Star began studying the emergence of information systems, and especially the internet, which was often presented as a new solution for communication issues. Together with Karen Ruhleder, Star studies how the introduction of an online collaboration platform affects collaboration in a community of geneticists (Star & Ruhleder, 1996). In this study, she combines her interest in science and collaboration with an interest in the design and use of information systems. Drawing on the work of Thomas Hughes, Star develops the idea that infrastructural changes are not just a matter of technicalities but involve social and organizational aspects as well. In the article, *Steps toward an ecology of infrastructure,* (Star & Ruhleder, 1996), she jointly proposes a research agenda for ethnographic studies of infrastructures, which she further develops in an article called, *Ethnography of infrastructures* (Star, 1999). Star and Ruhelder propose that infrastructures have certain

properties, such as being embedded within other structures, social arrangements, and technologies, being built on an installed base, inheriting their limitations and strengths from this base, and becoming visible upon breakdown (Star & Ruhleder, 1996, p. 113). Star furthermore suggests some 'tricks' for 'reading' and 'unfreezing' information infrastructures, such as identifying master narratives and their others, and surfacing invisible work (Star, 1999, p. 384). Star proposes a study of information infrastructures as both ecological and relational. Rather than viewing an information infrastructure as a thing with fixed properties, she suggests studying the wide range of effects and meanings it has for different people and users. The study of infrastructures has become a key orienting theme for on-going research within the STS community (Slota & Bowker, 2017).

3.2.3 Lucy Suchman – The reconfiguration of socio-material assemblages

Finally, this dissertation draws theoretical inspiration from the research of Lucy Suchman. Suchman's academic background is in critical anthropology and ethnomethodology. Like Star, Suchman has been a central scholar in STS, and their works have many commonalities. She has not only contributed to sociological studies of technology, but also fostered conversations within software development, in particular within artificial intelligence (Vera & Simon, 1993; Winograd, 1993). Suchman began her academic career doing research in Xerox Paolo Alto Research Center where she studied the attempts of computer scientists to design intelligent interactive interfaces for photo-copiers (Suchman, 1987). Suchman was interested in what it required to establish a successful interaction and relationship between humans and machines, a research theme which has remained central to her work.

One of Suchman's main critiques of the intelligent systems of the 1980's is that they were often based on the assumption that plans determine actions. Instead, Suchman argues that plans should be seen as resources for situated action. She gives the lucid example of how one would approach a series of rapids in a canoe.

> The plan might go something like 'I'll get as far over to the left as possible, try to make it between those two large rocks, then backferry hard to the right to make it around that next bunch.'… When it really comes down to the details of responding to currents and handling a canoe, you effectively abandon the plan and fall back on whatever embodied skills are available to you. (Suchman, 2007, p. 72)

Suchman also argues that intelligence should not be viewed as a purely cognitive phenomena, but has to be understood in relation to its interaction with the material and social world.

In studying the relationship between humans and machines, Suchman emphasizes the importance of paying attention to how work is represented (Suchman, 1995). In her article, *Making work visible,* she argues that representations of work are intertwined with broader organizational themes related to creating and maintaining invisibilities and the distancing and separation of different kinds of work. Making representations is not something reserved for ethnographers; it is part and parcel of working in an organization. Everyone who speaks and writes about work is involved in making representations, which are shaped by their own social position and values:

> The representations ethnographers create, accordingly, are as much a reflection of their own cultural positioning as they are descriptions of the positioning of others. This is not a problem or limitation to be overcome; it is a fundamental aspect of representational work, to be understood and incorporated into our practices and into what we produce. (Suchman, 1995, p. 63)

In the article, *Making a Case: Knowledge and Routine Work in Document Practices* (Suchman, 2000)*,* Suchman investigates knowledge work and routine work as two representations of work that play a role in the law firm she studies. She argues that with the rise of information as a dominant commodity, an appraisal of knowledge work has provoked a systematic deletion of other types of work. Suchman sets out to disrupt the dualism between knowledge work and routine work. She writes that knowledge work is often considered the work related to the manipulation and production of information, whereas routine work is related to the work which 'still' relies heavily on human interventions (Suchman 2000, p. 29). She argues that the separation of work into these categories misrepresents how work is actually practiced at the law firm: "The distinction orders not so much ways of knowing or acts of reasoning as it does identities, actors and distributions of material and symbolic reward among them" (Suchman, 2000, p. 43).

In Suchman's later work, she develops the notion of reconfiguration (Suchman, 2007, 2012). With this concept, Suchman wants to re-think, "configurations of the human and the machine," and she argues for, "articulating the differences within particular human-machine configurations, expanding our unit of analysis to include extended networks of social and material production and recognizing the agencies, and attendant responsibilities, involved in the inevitable cuts through which bounded socio-material entities are made" (Suchamn, 2007, p.6f).

Suchman uses the concept of reconfiguration to analyze how boundaries and responsibilities shift and are negotiated discursively as well as materially. According to Suchman, configuration is, "a tool to think with about the work of drawing the boundaries that reflexively delineate technological objects, and as a conceptual frame for recovering the heterogeneous relations that technologies fold together" (Suchman, 2012, p. 48). The notion of configuration is made up of two words: configure and figuration. To configure something refers to the act of arranging and ordering its elements in a particular form for specific purposes. Whereas the figure highlights the semiotic aspects of changes in language and signs that set expectations, goals, and motivations: "The effects of figuration are political in the sense that the specific discourses, images and normativities that inform practices of figuration can work either to reinscribe existing social orderings or to challenge them" (Suchman, 2012, p. 227).

3.3 Summary – The reconfiguration of work situations

The three researchers presented in this chapter suggest that researchers and practitioners should be attentive to the continuous interplay between work, information technologies, and organization. Where Zuboff draws on and connects to historical sociological conceptualizations and debates about work, Star and Suchman are more interested in what counts as work, and how work is represented in particular work situations. Star and Suchman shift our attention from sociological theories of organization and work to an interest in the language and worlds of the informants. In Zuboffs analysis of how IT affects workplaces, she provides some clear theoretical focus points such as the effect IT has on organizational structures and organizational relationships of authority. She sketches how the introduction of IT in organizations is involved in some general organizational shifts, such as the ways in which action-

centered forms of work are being displaced and replaced by intellective forms of work – often taking place on a computer. In the works of both Star and Suchman, we see some more post-structuralist tendencies. Star, for example, highlights the differences between how authority relations and organizational structures are described and made visible in discourses, and how they operate in practice; she poses that there is a difference between the alleged visible structures and the often invisible work they purport to describe. Suchman argues that the separation of work into that work which requires knowledge and that which requires the use of the body (what Zuboff calls the division between action-centred and intellective skills) is not just an 'innocent' description of work practices, but can also be used as a way for organizations to differentiate between workers, for example to distribute symbolic and material rewards (Suchman, 2000). In spite of these overall differences in research approaches, Zuboff, Star, and Suchman share many similar themes and concerns about the effects of IT on organizations and its employees.

Studying reconfigurations in the DCTA, I am guided by several of these concepts. In my research, I study three work situations, where public employees are working with data-driven technologies and visions. The information technologies in this study are related to self-service platforms, machine learning algorithms, and data-sharing infrastructures. The dissertation sets out to discuss organizational consequences of the current work going on around these data infrastructures, and to explore the challenges public employees face when working with or 'around' them. With the notion of reconfiguration, I highlight an attentiveness towards the representations and figures that public employees express and work-with, such as the no-touch customer, the customs officers' nose, bringing data home, and the data-driven organization. These emic figures provided an opening into the political and organizational issues that surrounded the work situations I studied. When I explore these figurations, I am attentive to relationships between visible and invisible work. What work is appreciated and highlighted, and what work gets relegated to the background, while still being important or even vital for the overall work situation to function. Finally, I am attentive to the redrawing of organizational boundaries and the distribution of organizational competencies and responsibilities.

4 METHOD

This chapter presents my methodological considerations and the methods used in the fieldwork. First, the chapter describes the construction of the data-driven vision as a research object. Second, it provides an overview of the data that have been collected, and the process of going from data collection to analysis. The chapter ends with a reflection on what the process of negotiating access with the tax administration communicates about the DCTA as an organization.

4.1 Constructing the research object – Three work situations

This dissertation explores how the DCTA becomes reconfigured through studies of work situations, where public employees work with new information technologies and technology-based visions. Methodologically, the study draws inspiration from the research approaches of actor-network theory (ANT) (Latour, 2005) and situational analysis (SA) (Clarke, 2005). These approaches are compatible with the theoretical orientation outlined in the previous section. They are empirically grounded approaches, which can be used in the description and analysis of how situations are composed by a multitude of human and non-human actors, which are co-constitutive of one another and the situation as such. ANT and SA suggest that researchers explore the complexities and ambiguities of social life, and argue for methods, "that can address and elucidate the complexities of situations as the grounds of social life (Clarke, 2005, p. xxix).

Studying information technologies and related organizational visions has its own challenges and complexities. In, *Researching partially existing objects* (Jensen, 2010), Casper Bruun Jensen reflects on his study of the development of an electronic patient record (EPR). Jensen frames the EPR as a partially existing object (Latour, 1996), an object in development, not yet stabilized. Jensen draws our attention to the fact that the EPR, especially in its implementation stages, is not a singular object but has a variable and changing ontology (Jensen, 2010).

> As one encounters 'it' empirically, the EPR is sometimes a word, a text, a vision, a procedure, a prototype, and interface and a database. One cannot decide in advance whether the referent is linguistic [or] rhetorical… Prior to empirical scrutiny, one simply cannot be sure whether the EPR is something 'envisioned' or something 'concrete'. (Jensen, 2010, p. 25)

The vision of the data-driven tax administration has emerged to me similar to what Jensen calls a partially existing object. Whereas the EPR was a partially existing technology, the data-driven tax administration might be described as a partially existing organization related to a range of partially existing technologies. As I began my ethnographic studies in the DCTA, I did so with the expectation that I would study specific technologies and encounter their materiality and effects as working technologies. But what I mainly encountered was people discussing in meeting rooms, writing and talking about the development and use of data-driven technologies and

visions. I learned that, "materiality is precisely what is lacking so much in technical projects" (Latour et al., 2011, p. 45).

The vision of a data-driven tax administration often appeared rather elusive, as illustrated in the prologue in chapter 1, when a manager suggested that there was a need to identify what the vision even meant. While it was elusive, it also appeared as a self-contained good, and most of the tax employees I spoke to knew of the vision and related their own work to the vision in various ways. I was unable to find much about the notion in the formal reports from the Ministry of Taxation, but once in a while it appeared in news articles, in job applications from the tax administration, and in internal PowerPoints. Metaphorically speaking, there was a 'frame' with the title of the data-driven tax administration, but the picture of the frame was missing. Throughout my fieldwork, I struggled to find this picture of the data-driven organization, and at several times I contemplated throwing the frame away. However, during the times when I had almost forgotten about the frame, it happened that an employee from the tax administration would suddenly connect their own work to the larger vision of the data-driven organization, reminding me about my initial research interest and making me reconsider it as the overall research object of my ethnographic study.

Encountering the data-driven tax administration

I arrived early in the DCTA to find a good seat in the canteen. Thirty minutes later the canteen was packed with tax employees, for some had arrived on busses from the neighboring tax departments. The Minister of Taxation was visiting the department for the first time. He was about to present the re-organization of the tax administration, 'From one to seven agencies' (Danish Ministry of Taxation, 2017). The Minister connected with the audience instantly. He told a story about one of his visits to another tax department. The tax employees had introduced him to an altar with candles, flowers, and a series of pictures showing former tax ministers. The tax employees told the Minister that this arrangement was 'the chamber of horrors' and warned the Minister that an empty frame was waiting in the drawer. Everyone in the audience laughed as the minister told the story. His second joke within the first 5 minutes.

The Minister's main purpose for the visit was to present the new organizational structure, but I paid particular notice as he mentioned the vision of the data-driven tax administration. He told us that he had initially been quite skeptical of this vision. As a liberal, he felt uneasy about a public organization collecting and storing large amounts of data on its citizens. But he had been convinced that becoming data-driven would make the tax administration more efficient and able to provide better services to the taxpayers.

It had been a while since I had heard someone speak about the data-driven tax administration, and it was the first time I heard it being expressed by someone outside the IT departments of the tax administration. While the vision had initially been central for my interest in the tax administration, it had largely faded to the background as I started to follow tax employees in their daily work. It was not something tax employees discussed on a daily basis. When the minister used the notion of the data-driven tax administration it thus came back to my attention. The minister did not, however, say much about the data-driven organization, and it seemed to me while everyone was applauding the term, no one seemed to be interested in elaborating what it means to be a data-driven organization.

Figure 4. Visit by the Minister of Taxation (Fieldstory, 9th of February 2018)

My research of the data-driven vision consists, on the one hand, of my conversations with tax employees and written materials that directly address the vision, and on the other hand, of the study of work situations, which I found to be related to the vision. With inspiration from situational analysis, I divide my fieldwork into three distinct work situations (see table 1). While there are overlaps between these situations, I consider them relatively distinct from one another. Each situation is concerned with specific departments, professions, and technologies. These three work situations form the basis of the three research papers presented in the second part of this dissertation. In the following, I introduce the three work situations with a focus on the main actors involved and the different locations I visited during the data collection.

In situation 1, I observed the work of frontline workers who talked with taxpayers over the telephone, social media platforms, and e-mails. This happened in a call center in Odense, which is located in Funen, the third largest island in Denmark. I also interviewed the frontline workers and managers working in the call center. I then visited one of the strategic headquarters of the tax administration located in Northern Copenhagen. Here I made three interviews, two of them with what I refer to as customer strategists – employees who were involved in thinking strategically about how the tax administration should relate to taxpayers. In this situation, I was

introduced to notions of 'channel strategies' and 'no-touch customers', which referred to ideal ways of being in contact with taxpayers. The research paper discusses how these ideals affected the work of frontline workers.

In situation 2, I relied on conversations with data scientists and the manager from the office for advanced data analysis, located at the DCTA's headquarter in Northern Copenhagen. I met with the office manager to plan and arrange the fieldwork, and interviewed several of the data scientists employed there, who also showed me around their offices and introduced me to their working environment. I then visited different parts of the customs department. I visited a monitoring unit in Southern Copenhagen, where two employees were monitoring the IT system that selected which packages to inspect for fraud. I also spoke to other employees who were involved in coming up with the rules that determined which packages to inspect for fraud. Finally, I visited the Copenhagen Airport, where I spoke to the customs officers working there, and I experienced a group of customs officers making an inspection of a truck.

In situation 3, I visited a wide range of places, but most of the time I was at the DCTA's headquarters in Northern Copenhagen. In this situation, I followed project participants working on an IT project to develop a new data-sharing infrastructure in the DCTA. I observed and spoke to a wide range of project participants, who had different roles and connections to the project. My main go-to person was the project manager, who was present at most of the meetings and situations that I visited. I also talked to many of the consultants who were guiding and helping the project along, as well as the customers of the project, who were employees in other public organizations. I went to four of the customers' organizations and interviewed them. In this situation, there were two research themes in particular that I became interested in. One was in agile methods, which was a new development method that the project participants often highlighted as a crucial feature of the IT project. The other was in the project's relation to a strategy of 'bringing data home', which was related to a wish to gain more internal control over the organization's data and to take on increased responsibility for the development of IT systems, rather than relying on private companies.

Table 1
Overview of work situations

Work situation	Technology in focus	Professions in focus	Primary locations visited
#1 Changing work practices of frontline employees	Self-service platforms, Telephone, Social media, E-mail	Frontline workers Managers Customer Strategists	Customer center (Odense, Funen) DCTA headquarter (Northern Copenhagen)
#2 Advanced algorithms in customs inspections.	Machine learning algorithms	Data scientists Customs employees Customs officers	DCTA headquarter (Northern Copenhagen) Customs monitoring unit (Southern Copenhagen) Copenhagen Airport
#3 Bringing home data infrastructures	Data sharing platform	Project participants (e.g., project manager, product owner, agile coach) Consultants Customers	DCTA headquarter (Northern Copenhagen) Events in Copenhagen

4.2 Overview: Data collection

The ethnography is based on an explorative research approach. Initially, I approached my ethnography in the DCTA with a broad interest in 'innovative data practices,' and how large data sets had become, "actors in the making of modern government institutions" (See Appendix A – Statement of project participation). With an overall interest in the vision of the data-driven tax administration, I began my inquiry in an open and exploratory manner. I wanted to observe and listen to tax employees to learn what issues and problematics they faced, and how the new technologies and visions affected their work. Coming up with research questions and themes has thus been an iterative process throughout the research. Following the formulation from Carse and Kneas, I see my research approach as one that aims to, "theorize from actor categories and empirical situations rather than philosophical first principles" (Carse & Kneas, 2019, p. 13). I further unfold this formulation in section 4.3, but first I provide an overview of the data collected for this dissertation.

I carried out fieldwork in the DCTA between June 2016 and December 2018. The most consistent period of my ethnography was between January and July 2018, where I visited the tax administration on forty distinct days. During the fieldwork, I collected qualitative data by observing, interviewing, and acquiring documents from different work situations in the tax administration. In the following, I briefly outline these different types of qualitative data, and how I collected and documented it. The data will be further presented in each of the research papers presented in part two of the dissertation.

4.2.1 Observations

My primary sources of data are observations within the tax administration. I use the term observation as a catchall-term for the various experiences I had in and around the DCTA. To provide a simple overview of these observations, I divide them into 70 distinct units. I define a unit of observation as an ethnographic experience that stood out as a distinct experience, which I have captured through fieldnotes under a common header. To exemplify this, I consider a workshop lasting 5 hours as one unit of observation, as well as a 30-minutes lunch with tax employees. This way of accounting for my observations also means that some experiences are left unaccounted. For example, I spent a significant time researching the digital infrastructures in one of the IT-projects, performing a form of digital ethnography (Pink et al., 2015). During a three-month research exchange, I 'checked in' on the IT project's communication platforms Jira and Confluence on a regular basis.

Table 2
Overview of observations

Type of observation	**Units**	**~ Hours**
Project meetings	34	58
Introduction to work practices	9	30.5
Events	4	9
Other	23	34
Total	**70**	**131.5**

In the table above I have divided my observations into four different types. I participated in 34 project meetings, where my primary role was to observe and listen to the tax employees talking about IT projects. At some of the meetings I would present myself at the beginning, and once in a while I was asked to give my opinion on a discussion. These project meetings involved the stakeholders of the IT projects I followed. The project meetings were varied in terms of their purposes. Some involved discussions about the needs of the customers and users of the IT solution, while others were status meetings where project participants discussed the progress of the IT project. All of the meetings occurred in offices at the tax administration. Another type of observation is what I have called introduction to work practices. Project meetings are obviously also a work practice, but I use the term here to highlight observations where I was introduced to a specific work practice of a public employee: For example, how a data scientist introduced me to his daily work, talking about the different software he used and his programming work, how I watched and listened to frontline workers talking to taxpayers through the telephone and writing to taxpayers through social media platforms, and how I participated in a customs inspection with customs officers inspecting a truck for possible tax fraud.

Furthermore, I participated in public events together with the tax employees. The events were a way for the tax employees to get inspiration about new approaches to the development of IT systems. An example of such an event was my participation in a network event for IT managers in the Danish public sector arranged by the Danish Agency for Digitization. Finally, the category 'other' represents more spontaneous moments which occurred due to my presence in the tax administration. These moments are conversation at lunch, at the coffee-machine, or in the corridors or offices of the tax administration.

I documented these different observations through fieldnotes. When I visited the tax administration, I carried a notebook along with me, and I adjusted my notetaking depending on the particular type of observation. At project meetings, I was able to write extensive fieldnotes as my participation in the meetings mostly was limited to observing. In writing fieldnotes, I first of all focused on the place/setting and the persons involved in the situation (Narayan, 2012). I noted down where my observations occurred and if there was anything interesting or peculiar about the

setting. I also noted which people were present and if they had a central role in the observation. Furthermore, many of my fieldnotes, especially from project meetings, were directly concerned with the discussions taking place. I noted down the different attitudes and opinions of the tax employees, often as close to verbatim as possible, while still being able to follow along in the conversation. At other occasions, especially in what I have called 'introduction to work practices', it wasn't always possible to write fieldnotes while observing and talking to the tax employees. At some of these encounters, I made sure to find a quiet location after the experience to write down my immediate thoughts and impressions. At times, I had an experience which I felt was particularly important, and it somehow stuck with me. In these instances, I often wrote more extensive field stories. These would either be later at the same day but could also be written weeks or months after the event. When writing field stories, my aim was to make a coherent and smooth narrative, and I tested out different ways of telling the story, focusing on different aspects of the encounter. These field stories have for example been used in the opening montage and in the example on page 33, and they are present as vignettes which I have used in each of the three research papers.

4.2.2 Interviews

During the fieldwork, I conducted twenty-one interviews: eighteen with tax employees and three with external customers. The interviews were semi-structured, and for each interview I prepared an interview guide (Creswell, 2007, p. 136; Hastrup et al., 2011, p. 78) to frame the conversations and to prepare the interviewees for the interview (see Appendix B – Examples of research protocols). Before the interviews, I sent the research protocol to the interviewee, in order to provide them with the overall topic and purpose of the interview. I also prepared for the interviews by making an extended list of questions for each of the themes outlined in the research protocol. I aimed for the interviews to be open-ended conversations, making it possible for the interviewee to elaborate on the topics and themes they deemed most important. The interviews were recorded and a majority were transcribed, some with help from research assistants.

Table 3
Overview of interviews

#	Interviewee	Date	Work situation		
			#1	#2	#3
1	Telephone agent	June 2016	▩		
2	Telephone agent	June 2016	▩		
3	Functional manager	June 2016	▩		
4	Call center manager	June 2016	▩		
5	Customer analyst	September 2016	▩		
6	Customer strategist	September 2016	▩	▩	
7	Self-service agent	October 2016	▩		
8	Data analyst	June 2017		▩	
9	Data analyst	July 2017		▩	
10	Project manager	August 2017		▩	▩
11	Software developer	February 2018		▩	
12	Information architect	July 2018			▩
13	Product owner	July 2018			▩
14	Project customer	July 2018			▩
15	Project manager	July 2018			▩
16	Scrum master	July 2018			▩
17	Senior user	July 2018			▩
18	Software developer	July 2018			▩
19	Project customer	July 2018			▩
20	Project customer	July 2018			▩
21	Customs officer	July 2018		▩	

4.2.3 Documents

During and after the fieldwork, I collected a wide range of documents, which have been valuable for contextualizing my observations and providing me with additional information about the three work situations. Some documents were acquired from employees in the tax administration, such as PowerPoint presentations describing departments and strategies of the tax administration or the PowerPoints which were

often used in the various project meetings. Other documents were publicly available, such as reports published from the Ministry of Taxation, the National Audit Office of Denmark, and the OECD.

Table 4
Overview of central documents

Documents	Examples	Publishers
Internal documents	PowerPoints & project documents	The DCTA
Strategies and Reports	Customer strategies & ministerial reports	Danish Ministry of Taxation OECD
Audit reports	Audits on IT systems or business areas	National Audit Office of Denmark
News Articles & Press releases	News about organizational changes and update on IT systems	Skatteministeriet.dk Computerworld.dk Version2.dk Dr.dk

4.3 From data collection to analysis - Follow the actors

In going from data collection to analysis, I have been guided by the popular notion from actor-network-theory to 'follow the actors' (Latour, 2005). In the process of my data collection, this refers first of all to speaking to tax employees about their work, and to participating in their daily work practices through observations and forms of shadowing (Czarniawska, 2008). It has been an explicit goal to listen to as many different voices and perspectives as possible in the different work situations I studied. As Susan Leigh Star suggests, it is the, "primary job of the sociologist: to listen to all the voices in a situation (as much as possible), including your own" (Star, 1991, p. 271). Second, and more substantially, I have used the dictum of following the actors to follow objects such as the data-driven vision, work situations, and figures. With a kind of recursive approach, I first followed the vision of the data-driven organization, locating it within concrete work situations in the tax administration. I then followed these work situations, unpacking the actors involved and the different goals and

challenges present in the work situations. Then, in each of the three work situations, I discovered and followed a figure (Haraway, 1997; Suchman, 2012), which especially guided me in moving from the work situations to analysis. I thus also 'followed the actors' in moving from observations to analysis.

This approach, which involved choosing between the wide range of ethnographic materials, often happened because I couldn't stop thinking about a particular encounter during my fieldwork. I experienced that some episodes simply got stuck in my mind. I think about these episodes as ethnographic moments (Strathern, 1999) and moments of disconcertment (Verran, 1999). For example, I remember being completely perplexed after visiting customs officers in the Copenhagen airport. I had visited the customs officers because of my interest in how the development of a new machine learning algorithm would affect their work. Before visiting the customs officers, I had mainly been talking to and visiting the new departments for data analysis and talked to the data scientists and project participants involved in the development of new algorithms. They had their own technical and organizational challenges, but I was completely struck by how very different the challenges which customs officers faced were. Although it was eventually going to assist the customs officers, I was made obviously aware about the fact that different social worlds have different concerns (Suchman, 1995). The difference and the detachment between these different types of work just seemed insurmountable.

The experience of ethnographic and disconcerting moments was at times directly related to a particular notion such as the experience I describe in bringing data home (see Research paper 3). To follow the actors, I have not just been interested in the discourses of the tax employees, but I also wanted to take discourses and make them central for my own analysis and writing. I was intent on following Latour's notion of catching up with the wild innovations of informants:

> In order to learn from them what the collective existence has become in their hands, which methods they have elaborated to make it fit together, which accounts could best define the new associations that they have been forced to establish. (Latour, 2005, p. 12)

Finally, the process from data to analysis involved a process of going back and forth between the empirical data and my own theoretical interests, informed by academic

debates and discussion. Like John Law describes, he experienced a moment where his field notes "started to produce signals" because, "data and theory interacted together in a way that resonated and amplified one another to produce pattern and repetition" (Law, 2004, p. 111). With this in mind, following the actors is not a denial of my own involvement in highlighting certain aspects of my data rather than others. This is also emphasized by Brit Winthereik, as she states that, "analysts as well as interlocutors both bring concepts into the ethnographic encounter" (Winthereik, 2020, p. 30).

4.4 Access reflections

Figure 5. Guest card for the DCTA (own photo)

Getting access is a well-known challenge in ethnographic studies (Neyland, 2008). In this chapter, I account for some complications I encountered in negotiating access with the DCTA and in studying IT projects and related work. I reflect on these complications and what they might tell us about the DCTA as an organization.

Without much experience in ethnographic studies prior to my dissertation, I learned the hard way the difficulties associated with getting access to conduct an organizational ethnography. It was by no means a straightforward task to get access to the DCTA and to study IT projects and related work. Prior to my fieldwork, a legal officer from the DCTA had signed a formal document agreeing to participate in the study and to provide access to conduct ethnographic fieldwork within the tax

administration (see Appendix A). During my studies, I was asked twice to sign a confidentiality agreement, making sure that I wouldn't share organizational details that could harm the organization economically. However, there was a long way from signing formal documents to actually getting access to study work in the organization.

When I reached out to employees in the tax administration, people were in general open and willing to talk to me. It wasn't difficult to set up interviews or to arrange one-day visits. The difficulties arose when I tried to establish longer-term arrangements that wouldn't require me to constantly reach out to the tax employee. This issue first became apparent in my studies of the risk score project (work situation #2). While the head of office had agreed to collaborate with me and suggested that I could follow the risk score project, he also told me that I would have to negotiate further access with the lead data scientist of the project. While the data scientist was open for interviews, and was happy to introduce me to his work, colleagues and the office in general, I still had to actively reach out to him to make new appointments. I was interested in a more permanent access to the data analysis office, which I felt was necessary to follow the risk score project more closely and learn about the work and culture of the data scientists. I talked to the data scientist about the possibility of following his work more closely, and if it was possible to get a working desk at the office, and to participate in their daily meetings. The data scientist seemed to be open at first, but he made it clear that he would have to talk to his superior, since he didn't have the authority to provide me with access to the office. After some time, the data scientist told me that it wasn't possible to get a more permanent access to the office - the office was full. At the same time, the data scientist told me that they were in fact about to close down the risk score project, and he suggested that I find another project for my research.

Reflecting on this experience, several things seemed to coincide in preventing me from getting a more regular access to the office. While I was told that the office was full, I also got the impression that they could not see the value in having me around at the office, while the data scientists were working on their computers. Furthermore, I learned that there is a risk in framing ones research around a project, and the development of an algorithm. If I instead had framed my research as studying how

the data-scientists worked, it might have been more understandable as to why I wanted more regular access to the office.

I ended up taking the advice of the data scientist, and found another IT project, which I could follow. In my studies of the risk score project, I came in contact with a project manager. He repeated to me that the risk score project was closing down but told me that he was starting up a new project, which I might be able to follow more closely. He told me that this IT project was much larger, and thus not as prone to being terminated as the risk score project was. The new IT project aimed to develop a data-sharing infrastructure to share tax data with external organizations. I was unsure if I should make this shift in research focus. I had already spent a considerable time getting acquainted with the data analysis office, I had been in contact with several of the employees working there, and I had studied academic literature on algorithms and artificial intelligence. The shift to a focus on the development of a new data-sharing platform, and to the management of larger IT projects in the state, required me to rethink my existing research plan. Since I had not managed to get a permanent or consistent form of access to the data analysis office, I decided to agree to the proposal of the project manager, who from the beginning of our conversations showed a significant willingness to include me in the project.

In studying the new data-sharing infrastructure (work situation #3), the project participants were much more open to include me in their daily work, than what I had experienced so far. I was quickly invited to participate in several project meetings, and to participate in a monthly advisory group meeting where the progress of the project was discussed with its many external stakeholders, including the customers of the project and several project advisors such as representatives from the Danish agency of Digitalization. However, while I was invited to more meetings, and to sit together with the project participants, I was still reliant on reaching out to the public employees. To enter the headquarters of the tax administration, one needs an ID card, and guests can receive a guest card. Whenever I visited the tax administration, I called or texted an employee from the project, often the project manager, who picked me up at the reception, where I had picked up a guest card. The project manager had told me that it should be possible to get a temporary access card so I could enter the building without having to reach out to the project participants. But every time I asked him

how far they were in the process, there seemed to be some kind of bureaucratic issue. After a project meeting, I talked with the project manager about the situation. The manager explained that he was still waiting on some of the administrative employees, to proceed with the case. One of the other project participants, who had heard our conversation pointed out how ironic it was that it was so difficult to provide access for researchers, when expensive consultants could become part of the system and get access almost instantly.

In 2015, the DCTA and the IT University of Copenhagen entered into an agreement of collaboration. The DCTA thus illustrated its willingness to collaborate with researchers and universities and appeared as a relatively open organization. However, as my examples from the fieldwork showed, the openness for research has its limits, and researchers have to be able to negotiate access beyond formal agreements. In my experience, it was clear how the tax administration wasn't as ready to collaborate with researchers as it was with consultants, who almost instantly were able to get temporary access cards and access to the internal IT systems. It seemed that my ability to get ethnographic access was completely dependent on my ability to negotiate with individual employees in the tax administration. The employees I met didn't have any official procedures to follow in order to provide me with access to the tax administration. I experienced a bureaucratic organization, where the tax employees I spoke to, such as the data scientist and the project manager, had to ask and make arrangements with other people in order to provide me with access.

The fact that the DCTA had clear procedures for working with consultants, and that it was more difficult for me as a researcher to get access to the tax administration, tells us something about what the organization is, and has been, geared towards. It showcases how work in the tax administration throughout the last decades has been supported by consultants, which has been common across the public sector in general. This might also tell us something about the perspectives which are dominant when a vision such as the data-driven organization is discussed internally in the DCTA. We might speculate that the dominant perspectives on the data-driven vision are related to questions of implementation rather than explorations of the potential issues and challenges; that the data-driven organization is seen as a solution rather than a problem to be explored.

Besides my difficulties to get access to the tax administration without relying on the continuous assistance of the tax employees, I also experienced some curious responses when I was about to communicate my research publicly. The 13th of March 2018, I was giving a public presentation based on my research and fieldwork in the DCTA. The event was arranged by the Data as Relation research project and was called Digital Dilemmas. The event had around 150 attendants mainly from public organizations in Denmark. My presentation was discussing the dilemma of whether public organizations should accept the use of machine learning algorithms when even developers had difficulties explaining how they worked. A week before the presentation, my supervisor received a phone call from one of the tax employees. He was concerned about what I was going to say in my presentation. I called the employee and talked him through my presentation, which made him more relaxed, and he didn't have any problems with my presentation.

Upon publishing the research paper, *Helping or intervening? Modes of ordering in public sector digitalization*, I reached out to the manager who had been my primary contact and had helped me arrange my observations and interviews. I wanted to let him know that I was about to publicize a research paper based on my fieldwork in his department. The manager redirected my research paper to the press department, instead of looking at the article himself. I was then called by an employer from the press department. The press employee mainly commented on more cosmetical details, such as making sure that I had used the right department names. But she then added that she was a bit troubled with the way that I was critical of the tax administration's work in dealing with taxpayers. I talked to her a bit about the argument, which I didn't find to be that critical myself, and then she just said that this area with digitalizing the relationship to taxpayers was actually one of the areas which the tax administration was really proud of.

These two examples highlight how the DCTA is conscious about what is being said and written about the organization. Tax employees are worried about the public image of the organization, and whether their work is being criticized. This awareness is understandable in relation to what the tax administration has gone through in the last two decades with one negative media story after the other, as I described in *section 2.2*

crisis. Many tax employees told me how the organization needed more positive stories about its digitalization work.

5 LITERATURE REVIEW: THE DATA-DRIVEN ORGANIZATION

This chapter discusses the recent interest in data as a research object. It provides a brief overview of recent academic literature concerned with the social, political, and organizational implications of intensified investments in and attention to data. The chapter is comprised of four sections. First, it presents different research approaches to understanding and conceptualizing data, and its wider societal implications. Second, the chapter more specifically addresses the notion of the data-driven organization, and the ideal of the data-driven organization is connected to the emergence of Big Tech. Third, it discusses the invisible work and secrecies surrounding the ideal data-driven organizations, highlighting that there is more to these organizations than efficient use of data and seemingly autonomous algorithms. Finally, the chapter addresses recent studies of how the vision of the data-driven organization influences public organizations.

5.1 Data as research object

The interest in data from academics and government institutions is not a new phenomenon (Rosenberg, 2013), but in the past decade there has been an increased push for data-driven forms of governance (Rieder & Simon, 2016). In 2012, Danah Boyd and Kate Crawford announced that the 'era of big data' had begun. They described big data as a cultural, technological, and scholarly phenomenon in need of critical questions (boyd & Crawford, 2012). In 2014, Rob Kitchin published The Data Revolution, calling for studies of, "the latest wave of information and communication technologies" (Kitchin, 2014). Where Boyd and Crawford's article mainly focused on big data in a social media context, Kitchin highlighted that the data revolution was a broader societal transformation affecting businesses, governments, and citizens.

Notions such as 'the era of big data' and 'the data revolution' imply that studying data means studying a particular moment in time, where data and a related set of technologies shapes the development of modern societies. A starting point for authors who argue that we live in an especially salient time for data studies is often the observation that information technologies, especially the internet, personal computers, and smartphones, have become an ever-present part of modern citizens' lives. When citizens use computers and smartphones, they generate data which can document, for example, how, when, and where the devices are being used. With new technologies for storing, sharing, and generating data, it is possible to share data at a higher velocity, to produce and store data in larger volumes, and to utilize a larger variety of data than ever before (Kitchin & Lauriault, 2014).

In researching the data moment, scholars have turned their attention to a 'plethora of terms' investigating the different technologies and visions of the data revolution: algorithms (McQuillan, 2016; Seaver 2017), databases (Ruppert, 2012; Dourish, 2017), smart cities (Madsen, 2018; Heaphy, 2019), data infrastructures (Aula, 2019; Bates et al., 2019), and data-driven organizations (Plesner & Justesen, 2020; Reutter & Spilker, 2019). Several researches have pointed out how data and the so-called data revolution constitutes a rather tricky research object. It is a tricky research object because the 'plethora of terms' are often ambiguous, highly connected and difficult to locate in practice.

In Michael Hockenhull's ethnography of tech-events, he describes the bombardment of data terms, which were used by presenters and participants at tech events. Hockenhull writes, "tech, whether it be data, digitalization, smart cities, or blockchain technology seems to be empirically related to a multiplicity of other technologies and actors through discourse, imaginaries, and material connections" (Hockenhull, 2020, p. 12). Hockenhull uses the notion of the hyper-relationality of data to communicate the entanglement of data technologies at a discursive level.

In Nick Seaver's study of algorithms, he writes that by studying algorithms he hoped to avoid hyped and vague terms such as big data and artificial intelligence. However, during his fieldwork, he found that even algorithms, which he thought would have appeared as concrete and easily definable entities, became elusive and difficult to pin down.

> As 'algorithm' drifted out of computer science and into popular and critical academic discourse, it seemed to signify a renewed concern for technical specificity. Where 'Big Data' was vague—originating in an overheated marketing discourse—algorithms were precise… Yet, just as critical scholars picked them up, algorithms seemed to break apart. (Seaver, 2017, p. 1)

Seaver discusses what he calls the terminological anxiety that surrounds data technologies and in his case algorithms. He dismisses the idea that the terminological anxiety is a shortcoming of social scientists who are failing to understand the technical details of their research objects. Instead, he argues that the terminological anxiety of algorithms is propelled by two factors: first, the discrepancy between how algorithms occur in textbooks, and how they occur 'in the wild' as part of organizations; second, the way in which social scientists study algorithms from an interdisciplinary approach, which aims to transgress and question disciplinary boundaries. In the attempt to transgress disciplinary boundaries, social researchers will unavoidably experience how different expert communities speak differently about algorithms and how they try to demarcate 'the proper way' of defining them (Seaver, 2017).

Several scholars have argued that the only response to the hype and terminological anxiety that surround the data revolution is to engage in empirical and critical studies. One of the attempts to reframe our understanding of data is critical data studies

(CDS). CDS also emphasizes the point that the data revolution and its plethora of terms have to be made more concrete and be scrutinized through empirical investigations: "By situating 'big data' technologies and data in contexts and thereby assessing its contingent, non-determinative role and impacts in society, critical data studies offer a less-hyped but more reasoned conceptualization of big data" (Dalton & Thatcher, 2014).

Rather than providing strict definitions or clearly defining the relationships between different technologies of the data revolution, several researchers have made the suggestion to study data as part of its wider processes of production, as socio-technical systems and as an inherently political actor. For example, researchers have proposed studying data value chains, which means paying attention to the different phases of the production and use of data (Löfgren & Webster, 2020; Flyverbom & Madsen, 2015). These authors argue that the production of data can be divided into different parts and stages similar to a manufacturing process. The data value chain proposed by Löfgren & Webster is: 1) collect data, 2) store data, 3) analyze & process data, and 4) utilize data – extract value (Löfgren & Webster, 2020, p. 4). The purpose of directing attention to the different stages of the production of data is to become more specific in general discussions of the characteristics of big data, for example to discuss how decisions in the early parts of the production process affect what is possible at later stages, or to compare different data value chains. Flyverbom and Madsen for example, show how different organizational contexts affect how data gets structured and used (Flyverbom & Madsen, 2015, p.157), and Löfgren & Webster highlight the fact that many different actors, both private and public, become involved in an entire value chain discussing the implications of multi-actor involvement (Löfgren & Webster, 2020).

Researchers have also framed their studies of the wider processes and relations of data as the study of data assemblages (Kitchin & Lauriault, 2014; Iliadis & Russo, 2016; Siles et al. 2020). In this view, data cannot be understood as a singular entity with a fixed identity but as shaped by the particular context in which it is used: "A data assemblage consists of more than the data system/infrastructure itself… to include all of the technological, political, social and economic apparatuses that frames their nature, operation and work" (Kitchin & Lauriault, 2014, p. 6).

Where studies of data value chains view the production of data as a linear and well-defined process, research into data assemblages highlights the multiple connections that shape and are being shaped by data. To study data as assemblages is to draw attention to how different worlds are being brought forth around data, and how existing relations and practices affect the emerging data practices.

Finally, researchers have suggested studying data politics, that is, "how data is generative of new power relations and politics at different and interconnected scales" (Bigo et al., 2019, p. 2). While the study of data politics shares many similarities with the idea of tracing and unpacking data assemblages, data politics highlights data as, "a language with a performative force" (Bigo et al., 2019, p.4), capable of reconfiguring relations between states, subject and citizens. With reference to Judith Butler, this term emphasizes the performative aspects of data, e.g. that whether or not data fulfills the hopes and desires it set out to, it still changes organizations as it directs their attention and investments in a certain direction and populates them with new organizational structures and materials.

5.2 The vision of a data-driven organization

The vision of the data-driven organization is often promoted in consultancy reports (Accenture, 2019), business magazines (McAfee & Brynjolfsson, 2012), and reports by international organizations (OECD, 2019; European Commission, 2020). These reports often highlight a small group of American companies: Apple, Amazon, Alphabet (Google), Microsoft, and Facebook as ideal data-driven organizations. These companies, also collectively referred to as Big Tech, have since the turn of the millennium made huge profits on selling data technologies and services. By 2020, Big Tech companies have become the most valuable companies in the world in terms of their market value.

The success of Big Tech is often explained by its unique ability to handle data. In the OECD report, *The path to becoming a data-driven public sector,* OECD presents Facebook as a role model for public organizations: "In the early 2000s, tech giants such as Facebook realized how digital platforms and the 24-7 connected citizen provided the ideal context to collect and reuse data for business purposes" (OECD, 2019, p. 24).

The business and management scholars McAfee & Brynjolfsson also argue that data-driven organizations simply outperform their competitors.

> Before long, [Amazon] developed algorithms to predict what books individual customers would like to read next—algorithms that performed better every time the customer responded to or ignored a recommendation. Traditional retailers simply couldn't access this kind of information, let alone act on it in a timely manner. It's no wonder that Amazon has put so many brick-and-mortar bookstores out of business. (McAfee & Brynjolfsson, 2012)

Big Tech companies handle an overwhelming number of daily users on their websites and platforms. In 2012, Google received over 3 billion daily searches, and Facebook 3 billion like actions per day (Kitchin, 2014). The companies store data in large distributed databases and use it to continuously change and update websites that are customized and personalized to their different users. Some of the companies even have famous algorithms such as Google's PageRank and Amazon's A9. The A9 algorithm ensures that Amazon's customers are offered the products they desire when they search on Amazon's websites, and the PageRank algorithm ranks and displays the 'most relevant' websites based on the search query of users. These algorithms are well-guarded trade-secrets, and objects of speculation and awe.

With access to large amounts of data and 'mythical algorithms' (Ziewitz, 2015), only the imagination seems to set the limit for the large tech companies. Besides using data to improve their existing services and products, the companies have frequently considered new ways to expand their relevance and develop new products and services. Two examples of this are Google's attempt to detect upcoming influenza epidemics (Ginsberg et al. 2009), and Facebook's psychological studies of its users' emotional states (Kramer et al. 2014).

The Google Flu Trends project, which ran between 2003-2008, was initiated by a group of Google employees. In an article published in Nature, the authors argued that they could predict the outbreak of influenza epidemics by relying on search queries from Google users. The reasoning was that if several users from the same geographical region were searching flu symptoms it indicated the emergence of a flu outbreak. With this method, employees from Google suggested that they could detect flu outbreaks

faster than traditional surveillance systems, such as the US Centers for Disease Control and Prediction (Ginsberg et al. 2009).

Another example is from 2014, where employees at Facebook published a research article in the Proceeding of the National Academy of Sciences (Kramer et al. 2014). The authors argued that they could show how Facebook users were affected by each other's emotional states, as expressed through their content of the Facebook social media platform. The authors of the article argued that they could influence Facebook users' emotions, by presenting them with particular content on their news feeds.

These two examples illustrate how Big Tech is relentlessly speculating about how they might utilize their data for new purposes and increase their organizational relevance. However, the examples have also prompted critical questions about the tech companies and their use of big data (Lazer et al., 2014). In 2013, The Google Flu Trends project failed greatly as it missed the peak of the flu season by 140 percent thereby turning, "the poster child of big data into the poster child of the foibles of big data" (Lazer & Kennedy, 2015). Furthermore, Facebook's emotional contagion experiment has been heavily criticized for its ethical implications, such as the limited agency and control users have in ensuring that their information is responsibly disclosed and shared (Selinger & Hartzog, 2016). While business magazines and Governmental reports often link the success of big tech to their ability to utilize data and algorithms, it is pertinent to ask critical questions both in terms of the ethics of their data usage, but also the accuracy and efficiency of their data analysis.

5.3 Secrecy and invisible work in data-driven organizations

In recent years, several studies have commented on the broader societal implications of data-driven organizations and provided more nuanced and detailed studies of their operations. A central theme has been the secrecies and invisibilities involved in the operations of these organizations.

Researchers have warned about the emergence of a black box society (Pasquale, 2015), the threat of algocracy (Danaher, 2016), and the rise of surveillance capitalism (Zuboff, 2019), arguing that democratic values such as personal freedom, freedom of speech and organizational transparency are at risk of being undermined and ignored

by data-driven organizations. These studies highlight the central role data and algorithms play in the way modern organizations operate. Algocracy is 'rule by algorithms', and John Danaher uses it to describe, "a system in which algorithms are used to collect, collate and organize the data upon which decisions are typically made" (Danaher, 2015, p. 3). The 'threat of algocracy' is that societies become more secretive and authoritarian as they come to rely on algorithms that are difficult to scrutinize and understand for the general public and for experts themselves.

Frank Pasquale characterizes different kinds of secrecy and opacity involved in organizations' use of data and algorithms. He distinguishes between technical and legal forms of secrecy. First, the algorithms used in many of the tech companies, are difficult to understand because of their technical nature (Seaver, 2014; Burell, 2016; Bucher, 2016), a trend which has been exacerbated by the recent push for more advanced algorithms, such as the use of machine-learning techniques. Second, Pasquale also notes that there are legal reasons for the secrecy of algorithms. Because the algorithms are the main competitive advantage for organizations, their source codes are protected by legal means. Pasquale suggests that the increased power of tech companies should lead to a push for putting them under more democratic scrutiny.

Examining the invisibilities of data-driven organizations is also impacted by the interrelations between various actors. Shoshana Zuboff emphasizes this in her analysis of the asymmetrical relationships of visibility between tech companies and their users. While Google and Facebook have access to all their users' information, their users do not know what data these companies have and how and when they use it. Like Pasquale, she highlights that the tech companies have an economic incentive to keep this information secret, since having access to their data is an essential part of their business models. She also emphasizes the economic incentives for tech companies to keep users hooked on their platforms and make them spend as much time as possible using their services, since this increases the platforms' relevance for advertisements, which are the biggest source of income for the companies. Zuboff argues that the tech companies exercise a new form of power, as they have 'colonized' people's most intimate and private sphere.

The framing of data-driven organizations as part of broader societal developments such as the black-box society, and the increase in surveillance capitalism both highlight a more problematic story of data-driven organizations. The secrecies and invisibilities that the studies emphasise is viewed as a consequence of the way organizations *use* data and algorithms. Another strand of research has focused on the invisibility involved in *producing* the data and algorithms and *making them work*. Where the former studies to some extent reproduce the narrative that the tech companies' success is related to their use of algorithms and data, the latter studies re-direct our attention to the human workers, who make it possible in the first place for the data-driven organizations to function.

Kjeld Schmidt writes that social researchers should be attentive to the broader work and supply chains involved in the operations of data-driven organizations (Schmidt, 2015). Schmidt draws attention towards the production chains of the tech industry. He points out how the production of iPhones are outsourced to Asian sweatshops, where workers sometimes work at, "the brink of suicide, working for up to four thousand hours per year" (Schmidt, 2015, p. 348).

Where Schmidt focuses on tangible products such as smartphones and computers, other scholars have drawn attention to the production process of the seemingly autonomous algorithms. These studies have unpacked how algorithms are supported by a wide range of activities, which are often ignored and left out of official discourses and documents, including workers who support the data collection or who reverse or ignore algorithmic decisions. With these studies, we become introduced to new organizational roles such as human quality raters (Bilic, 2016), data janitors (Zuboff, 2019), content moderators, and call center operators (Irani, 2015a).

Pasko Bilic notes how although Google presents itself as a neutral and objective provider of information, in practice it cannot provide "purely technical solutions" and has to rely on the, "continuous tinkering and tweaking of its search algorithm" (Bilic, 2016, p. 1). Based on leaked business documents, Bilic discusses the role of human quality raters, who have worked to find and fight spam on search results since 2004 (Bilic, 2016, p.4). The main objective of human quality workers is to rate Google's search results and assess their quality. Bilic notes how, "it is not fully clear how the

paid work of human raters impacts the Google search algorithm and, more importantly, if it affects page rank" (Bilic, 2016, p. 5). The paper by Bilic highlights how Google relies on work relations that are opaque for outside observers, making it difficult to assess the extent to which more manual forms of work are involved in sustaining and supporting algorithms and their decisions.

Lilly Irani has also drawn attention to the more manual forms of work involved in data-driven organizations. Like Bilic, she argues that companies such as Google, Amazon and Facebook hide the, "delivery people, stockroom workers, content moderators, and call center operators," in order to produce an "automated experience" (Irani, 2015a). She also argues that these companies, "rely on low-status workers' smarts to power the companies' seemingly miraculous algorithms and information systems" (Irani, 2015a). Irani challenges the narrative that the success of data-driven organizations can be explained by appealing to autonomous algorithms. She is interested in how work becomes displaced, including how responsibilities and competencies become allocated as algorithms are introduced into work situations.

One of Irani's key examples comes from her study of the Amazon service known as Amazon's Mechanical Turk (AMT) (Irani & Silberman, 2013; Irani, 2015b). The service is named after a seemingly self-playing chess board invented in the 18th-century. The chess board was a trick, which worked by hiding a chessmaster inside a case underneath the chessboard, controlling the pieces with a magnetic linkage (Irani, 2015c, p. 730). Amazon used this story to promote its online platform AMT that worked to delegate an overall work task to a wide array of workers, each working on a simple sub-task of the wider system. The separation of a work task into many small components, which are then delegated to many workers online, has been termed micro-work. Several scholars have noted that micro-work has so far played a significant role in supporting the development and operation of applications associated with artificial intelligence (Irani, 2015c; Tubaro et al. 2020). Micro-work is for example used when training or preparing machine learning algorithms, which relies on large datasets of a wide range of images that have been labelled and tagged according to the specific needs of the algorithm.

The use of micro-work has several organizational implications. One such implication is the way that attention to micro-work questions the perceived efficiency of applications of artificial intelligence. It shows that the development of advanced algorithms are in fact highly labor-intensive, rather than alleviating the need for workers: "Because of the wide scope of application of micro-work, it is a structural component of contemporary artificial intelligence production processes – not an ephemeral form of support that may vanish once the technology reaches maturity stage" (Tubaro et al. 2020, p. 1). Second, the ways micro-work has been implemented has allowed organizations to outsource work, thereby relinquishing their responsibility to the workers involved in performing work tasks.

The secrecies of data-driven organizations and the invisible work involved in operating and supporting data-driven technologies, also have large implications for public organizations looking to adopt more data-driven forms of governance. Public organizations operate under different conditions than private companies, and cannot in the same way hide their operations, nor rely on low-paid workers whose responsibility they have outsourced to external organizations. This leads to the questions, if data-driven practices are in fact highly labor-intensive, how do public organizations deal with allocating responsibility between their different workers, and do they have the resources to support their data infrastructures?

Several researchers have called for more detailed studies of how work in public organizations are affected by the data-revolution and the recent wave of digitalization (Plesner et al. 2018; Winthereik, 2018; Møller et al., 2020). Brit Winthereik (2018) has proposed studying data work as, "all the things that goes into producing the condition of possibility for what we typically think of as data work like data entry, analytics, etc." (Winthereik, 2018, p. 15). With this definition of data work, Winthereik broadens the notion of the work involved in data-driven organizations, and asks researchers to focus on the backstage work, which is part of the data-driven organization but often left out of discussions on data work.

5.4 The data-driven vision in public organizations

As already mentioned, the economic success of big tech coupled with their abilities to store and use large amounts of data have made them organizational role models

praised by international organizations such as the OECD and the European Union. According to these governmental bodies, public organizations should learn from big tech and become more data-driven. Following the recommendations from EU and OECD, many governments have invested in data infrastructures and developed national strategies on data and artificial intelligence (The Danish Government, 2019; Van Roy, 2020). However, the governmental reports also note that becoming data-driven is not necessarily a straightforward task for public organizations.

> Organisations from the public and private sector often face legacy challenges inherited from analogue business models, ranging from outdated data infrastructures and data silos to skill gaps, regulatory barriers, the lack of leadership and accountability, and an organizational culture which is not prone to digital innovation and change. (OECD, 2019, p. 24)

In recent decades, public organizations have experimented with using data and algorithms to support a wide range of different tasks. Health care organizations are looking to create more personalized health-services through data-driven technologies (Høyer, 2019), and to develop new data infrastructures that will foster a more seamless exchange of data, with digital solutions that are easy to use (Wadmann & Hoyer, 2018). Security agencies attempt to use data mining to fight the war on terror (Zarsky, 2011), and establish what has been called algorithmic security (Amoore, 2009). Police departments experiment with predictive policing, where advanced algorithms are used to predict where crimes are likely to occur, and to manage and allocate their resources (Benbouzid, 2019). Municipalities and Government agencies imagine how data-generating sensors will create smart cities improving the life of citizens and securing a more efficient use of resources (Madsen, 2018; Löfgren & Webster, 2020). Not to forget tax administrations such as the IRS, who experiment with algorithms to detect those taxpayers who are most likely to file false tax returns (Zarsky, 2013). These examples illustrate how public organizations expect data and algorithms to create more efficient organizations that are able to make more accurate and evidence-based decisions.

Several scholars have discussed how the interrelation between public and private actors is affected by public organizations' pursuit for increased efficiency and accuracy. Researchers have argued that public and private organizations become intertwined in new and complex data relationships, where the boundary between

public and private actors become blurred (Pasquale, 2015; Löfgren & Webster, 2020; Tupasela et al. 2020). These new and complex relationships arise as new forms of expertise and software are introduced in the different stages of a government's data value chains (Löfgren & Webster, 2020). This is supplemented by the rise of national data ecosystems, where public and private institutions converge (Tupasela et al. 2020). Rob Kitchin writes that governments become involved in "technological lock-ins", which create, "a corporate path dependency that cannot easily be undone or diverted" (Kitchin, 2014, p. 182).

Many researchers argue that public organizations' increased attention to data, and their investments in data infrastructures, should be seen as an intensification of neo-liberal policies. Dan McQuillan writes that, "the dialectic that drives AI into the heart of the system is the contradiction of societies that are data rich but subject to austerity" (McQuillan, 2018a). According to Jo Bates, open government initiatives in the UK can be seen as an intensification of neo-liberal policies, where private companies take on tasks previously performed by public organizations, and furthermore that public investments favor the growth and development of private companies. She also argues that the creation of data infrastructures to openly and freely share government data can be seen as a marketization of public services supported by taxpayers' money (Bates, 2014).

On the other hand, we also hear how governments attempt to seize the data moment and secure its financing by making use of public data and by making lucrative business deals with private actors. Several researchers have highlighted how the Nordic countries are reimagining their welfare states in the image of the data revolution. Tupasela et al. (2020) shows how Denmark and Finland, "view the accumulation of data as a way of partially financing the welfare state"(Tupasela et al., 2020, p. 9), and Maguire & Winthereik show how municipalities enter into agreements with Facebook and Google to build data centers, which in their construction and operation are imagined to foster new jobs (Maguire & Winthereik, 2019). Furthermore, Liam Heaphy has argued that a data-driven rationale leads, "pressure organizations to take control of their own data" (Heaphy, 2019, p.1), a trend which might cause public organizations to take on more responsibility for its services, rather than handing it over to private actors. These studies highlight that the investments in data and

algorithms are unavoidably linked to political questions, for example about how to distribute the responsibility for data infrastructures between public and private organizations.

The studies mentioned so far in this section discuss various consequences of the experiments and investments in new data practices. Ethnographic studies have addressed the vision of the data-driven organization more directly, and discussed the challenges public employees face when working with and towards this vision. A finding that reoccurs is that the vision is elusive and is used for speculation and experimentation in public organizations. Michael Hockenhull and Marisa Cohn use the example from a discussion at a Danish municipality around an almost empty PowerPoint slide with the title, "Data-driven what do we mean?" (Hockenhull & Cohn, 2021). They discuss how a Danish municipality engages in speculative data work, where employees debate how the future of the organization can be linked to data. In another study, Ursula Plesner and Lise Justesen also find that the data-driven vision is quite elusive and linked to speculation.

> In our initial fieldnotes we highlighted that there seemed to be a concern with becoming more data-driven, but also that management had little idea about what the outcome would be yet, as both the idea of working with data exploratively and the idea of hiring employees with data skills were new at the center. (Plesner & Justesen, 2020, p. 268)

In their article, Plesner & Justesen point to the tensions between the imagined possibilities of the data-driven organizations, particularly data-driven management, and the lack of more concrete goals in the projects they followed. Finally, Lise Reutter and Hendrik Spilker also discuss what they call the fluid and diffuse goal of becoming data-driven. In their study of the Norwegian Labor and Welfare Administration (NLWA), Reutter & Spilker introduce us to the work involved in developing machine learning algorithms. They find that the challenges of NLWA in using machine learning algorithms are, "if not insurmountable, at least far larger and more demanding than expected" (Reutter & Spilker, 2019, p. 104). They highlight obstacles associated with getting access to data, with the quality of the data that currently exists, and complex legal issues.

In these ethnographic studies of the data-driven vision, we become aware that although the rhetoric of the data revolution has become popular in public organizations, the vision might not be as developed internally where organizations might in fact struggle to make use of data and related technologies. While experiments with new data technologies and infrastructures might not turn out according to expectations, they are still involved in changing and altering relations between governments, public employees, citizens, and private organizations.

6 CONCLUSION

The final chapter presents the conclusions of the dissertation. First, it distinguishes between studying information technologies from an implementation perspective, from a critical perspective, and from the study of reconfigurations. Second, it provides a summary of each of the three research papers including their findings and contributions. The third section discusses some cross-cutting themes throughout the three research papers, providing reflections on the data-driven visions effect on public organizations.

6.1 Reconfiguring work situations

Taken literally, the vision of the data-driven administration alludes to a naïve technological determinism, as it stresses data as a key organizational driver and taps into popular ideas such as the automated workplace (Wajcman, 2017). What seems othered and left out of this catchy vision are the various workers who are supposed to manage, implement, support, and maintain the emerging data infrastructures. Discussion around the current hype and interest in data-driven technologies and visions can easily become crystalized into two extremes. On the one hand, scholars question how new information technologies and related visions can be realized and implemented in order to provide the increased organizational efficiency and quality of services they promise. In this view, the data-driven organization is the next progressive step on the pathway of modernization. This approach is exemplified, for instance, in discussions of government maturity models; stepwise models which present different levels of digital maturity in organizations and are used prescriptively to suggest how organizations should move through a series of changes in order to achieve a digital transformation (Veit & Huntgeburth, 2014; Jæger, 2020, p. 53). Another automatic response is to criticize how data-driven technologies and visions constitute a deprived view of social interaction and organization, and leads to the enhancement of authoritarian and undemocratic societies. Some exemplary critical studies include Zuboff's study of surveillance capitalism (Zuboff, 2019), and Dan McQuillan's view of AI and data science as a "form of neo-platonism" (McQuillan, 2018b).

The purpose here is not to dismiss studies of best practices for implementation or critical studies of data technologies and visions, I draw on several of these in my dissertation. However, I have attempted to utilize a more cautious approach, investigating what role data-driven technologies and visions play in specific work situations. Rather than accepting technology-based visions at face-value, I explored how these technologies and visions become translated and adapted to local circumstances. As the technologies and visions are adapted and incorporated into work situations, how do they reconfigure organizational relationships and redistribute competencies and responsibilities? The three research papers in this dissertation are all examples of how work situations are reconfigured as public employees work with data infrastructures.

6.2 Overview of research papers - Figures and findings

Paper 1: Helping or intervening? Modes of ordering in public sector digitalization

Short summary: The first paper investigates how frontline workers deal with new imperatives of digitalization. It focuses on the frontline workers employed to assist taxpayers with understanding, delivering and filling out tax documents. Since the organizational restructuring of DCTA in 2005, the role of frontline workers has changed significantly. Before the reform, frontline workers had a primary role in receiving tax documents from taxpayers and helping them with face-to-face guidance. After the reform, the self-service platform TastSelv became the primary place for taxpayers to hand-in and fill out their tax documents. With the taxpayer ideal of the no-touch customer, and channel strategies emphasizing online communication channels, the frontline workers have to some extent become secondary and not appreciated in the formal strategies. This paper investigates how this shift affects the work of frontline workers.

Findings: The paper finds that frontline workers still play a crucial role at 'the margins of data-infrastructures'. Frontline workers, who are available on telephones, e-mail and social media, ensure that taxpayers can 'deliver data' at the self-service platforms of the tax administration. While frontline workers are recommended to guide and teach taxpayers how to fill out their tax documents online, this paper finds that frontline workers at times have to leave aside this ideal. Instead, they actively help taxpayers fill out their digital tax documents. As they do so, they themselves become aware that they are doing "more than they should", indicating that they are uncomfortable with helping out taxpayers, rather than being pleased and satisfied with providing help to those taxpayers who need it.

Contributions: The paper contributes to current debates about how digitalization affects the work of frontline workers in public organizations (Bovens & Zouridis, 2002; Pors, 2015; Buffat, 2015; Schou, 2018). It contributes to ethnographic studies of organizational change by combining insights from STS with studies on the digitalization of public administrations. With the concept of modes of ordering (Law, 1994), the article emphasizes the value of a non-linear approach to change, and of

instead directing attention to the various translations that occur as digital reforms turn into organizational strategies and organizational practices. With a focus on frontline workers, we see that the effort to turn taxpayers into no-touch customers is a mode of ordering that is contested in practice. The paper furthermore emphasizes that discretion is not just an individual capacity, but a response shaped by the material and social elements of a situation.

Paper 2: Organizing artificial intelligence – Representing work in the Danish Customs and Tax Administration

Short summary: The second paper investigates organizational challenges related to the implementation of a machine learning algorithm. The DCTA's customs department has been heavily criticized by the National Audit Office of Denmark in their work to control and inspect imported goods and packages for tax fraud (National Audit Office of Denmark, 2018). In response to this criticism, the DCTA suggested developing a new machine learning algorithm to detect fraud based on historical data. This paper explores the work to develop a new machine learning algorithm from the perspective of the project's three different stakeholders: the data scientists, who work to develop a new algorithm; the system monitors, who manage and monitor the existing IT-system used to select packages for inspection; and the customs officers, who work to inspect packages for fraud. The paper explores the representation or figure of the customs officer's nose, which reminds the project participants that the new machine learning algorithm eventually has to fit in with the customs officer's work.

Findings: The paper finds that the development of machine learning algorithms are likely to increase the distances between existing organizational units, as their development relies on new forms of expertise. Furthermore, since machine learning algorithms are generated from data rather than by imitating the work of experts, which was the case with earlier forms of artificial intelligence, their development might direct attention towards existing IT-systems, at the expense of the work of the people who are eventually going to act based on their decisions. This can lead the development of machine learning algorithms to be decontextualized and to misalign with the work situation they are supposed to support.

Contributions: The paper contributes to critical studies of algorithms (Seaver, 2017; Kitchin, 2017; McQuillan, 2018b). Much research in the critical studies of algorithms has focused on private companies, the effects of algorithms 'at work', and on the black boxed nature of new algorithms. This case contributes with an empirical case from a public sector organization and provides an example of the failure to implement a machine learning algorithm. Failure stories are important to remind us about the difficulties of making advanced algorithms work in practice, and to challenge the view that they are already ever-present in contemporary organizations. The theoretical contribution of this paper is to connect earlier STS studies of the organizational implications of artificial intelligence (Star 1995b; Suchman 2007; Forsythe, 2001) with today's attempts to implement artificial intelligence in the form of machine learning algorithms. In the 1980's, STS scholars argued that expert systems were likely to be decontextualized because they assumed that the experts' work could be made explicit. This paper finds that machine learning algorithms are also likely to be decontextualized but due to the reason that their development does not draw attention to experts' work but rather to the data that are produced in a work situation.

Paper 3: Bringing data home – The reconfiguration of public data infrastructures

Short summary: The third paper investigates the attempt of the DCTA to take back control and responsibility of data infrastructures from private suppliers. It studies an IT-project with the aim of developing a new data-sharing infrastructure for sharing tax data to external organizations such as municipalities, unemployment funds, and pension funds. Prior to the new data-sharing infrastructure, the sharing of tax data was outsourced to a private IT supplier. That arrangement was considered to be expensive for the recipients of tax data and located the control and responsibility of data sharing to a private supplier. In the paper, I explore the figure of bringing home, which the project participants used both in relation to bringing home the control over data but also to bringing home the competencies for managing and developing data infrastructures.

Findings: This paper finds that public employees orient their work to develop new IT-systems around the figure of 'bringing home'. At a discursive level, the figure of

bringing home signals an attempt to take back the responsibility for developing and maintaining public IT systems from private actors. The figure suggests that a rather smooth shift is possible, and that there exists a clear boundary between the 'home' of the public sector and its 'away' of the private sector. In the development of a new data-sharing infrastructure the paper shows how 'bringing home' is translated on a technical and an organizational level. At the technical level, bringing data home indicates a separation between the underlying IT systems producing data, and the data itself. By bringing data home, the DCTA is still dependent on its private suppliers who are responsible for the underlying IT systems. At an organizational level, the DCTA still has to include consultants and external developers in the development of new data infrastructures, as project participants face challenges associated with the 'maturity' of the tax administration. Bringing home IT competencies also means to bring in consultants and external developers to work within the offices of the tax administration, and alongside the tax employees.

Contributions: The paper participates in recent debates on data infrastructures (Bates et al., 2019; Edwards, 2010) and contributes to discussions about the growing privatization of public services and infrastructures (Bates, 2012; Tupasela et al., 2020; Dunleavy et al., 2006). It contributes with an empirical case of the development phase of a public data-sharing infrastructure, providing insights into the negotiations and discussions that happen in these initial phases. Several studies have highlighted that the uptake of data technologies in the government are involved in a blurring between public and private boundaries (Tréguer, 2019; Löfgren & Webster, 2020), and that public organizations become dependent and 'locked-into' private solutions (Kitchin, 2014). This paper attempts to qualify and discuss these claims and highlights the on-going negotiations and shifting associations between public and private actors. It illustrates that a clear distinction between public and private actors might work on a discursive level but shows how the redistribution of responsibility and control happens at many layers simultaneously.

Table 5

Overview of research papers

Research paper	Figure	Organizational problem	Organizational solution	Configuration
1# Helping or Intervening? Modes of ordering in public sector digitalization	The no-touch customer	Taxpayers who need help by frontline workers to fill out tax documents are expensive.	Frontline workers should guide and educate those taxpayers who contact the tax administration. Taxpayers should be made able to solve their future tax-related issues on their own.	Taxpayers cannot always be guided and educated. In some situations, Frontline workers help taxpayers directly by filling out their tax documents. Frontline workers are embarrassed about 'doing more than they should', which highlights their work as somehow depreciated and unrecognized, while it is still an essential part of making sure taxpayers can deliver data to the self-service platforms.
2# Organizing artificial intelligence – Representing work in the Danish Customs and Tax Administration	The customs officer's nose	Detecting tax fraud on packages imported to Denmark are inefficient, and current rule-based algorithms don't work properly.	Machine learning algorithms can provide a more accurate detection of fraud than existing rule-based algorithms.	Customs officers currently lack the proper resources to perform efficient inspections of fraud, regardless of the method used for selecting packages. The organizational distance increases between those who are supposed to act, and those who decide when and how to act.
3# Bringing data home: The reconfiguration of public-private boundaries	Bringing data home	The DCTA is dependent on private actors to share tax data – an arrangement that is expensive and locates knowledge about data sharing outside the tax administration.	To bring home data and IT competencies. To take back responsibility for how data are shared externally, and for developing and managing new data infrastructures.	The central IT systems are owned by private companies, and the tax administration has to replicate the data from the private companies in order to get control over the data. The tax administration relies on expertise from consultants to develop new IT infrastructures as they lack this expertise.

6.3 Cross-cutting themes

Each of the three research papers, which are presented in their full length in part II of the dissertation, provides insights into how work situations in the public sector become reconfigured around the implementation and use of information technologies. The analysis of the work situations show how organizational boundaries are re-drawn both internally between public employees, and externally to taxpayers and private actors. The reconfiguration includes a re-distribution of competencies and responsibilities and the re-drawing of boundaries between visible and invisible work. In the following, I draw out some of the overall and cross-cutting themes of the three research papers. In this way, I answer the research question *how work situations are reconfigured as the Danish Customs and Tax Administration strives to become data-driven.*

First, the recent focus on data in the public sector has led to an increased specialization of work around information infrastructures. As public organizations experiment with data technologies, they invest in new departments and personnel and they hire consultants and external developers. In the last decade, the DCTA has invested heavily in new offices and expertise related to IT and data: offices for advanced data analysis, data governance, and agile methods, new professions such as data scientists, data engineers, agile coaches, and digital strategists have become part of the tax administration. This is a trend that has culminated with the creation of the IT and Development Agency in 2018. My articles suggest that the current investments in information technologies, such as machine learning algorithms and data-sharing infrastructures, are extensions of existing information infrastructures and involve the redistribution of competencies and responsibilities in work situations. The paper, *Organizing Artificial Intelligence,* shows how the development of machine learning algorithms extends and reproduces existing organizational boundaries. The paper highlights how data scientists constitute a further specialization of work in a situation where the responsibilities of a customs inspection have already become divided between those who inspect packages (customs officers), and those who decide which packages to inspect (system monitors). With the potential use of machine learning algorithms, an additional specialization occurs in the layer of those who decide which packages to inspect (data scientists). In the paper *Bringing data home,* I highlight how existing competencies become redistributed as the tax administration seeks to take back some responsibility for sharing tax data. By developing a new data-sharing

infrastructure, tax employees will become responsible for making data-sharing agreements and get an overview of the tax administration's data-sharing. While the DCTA does take back some of the responsibilities and competencies related to the sharing of tax data, they also reproduce some of the existing dependencies and create new ones. First, private suppliers are still responsible for the underlying IT systems generating the data. Second, the new data-sharing infrastructure is developed by external developers, making it an open question to which extent the tax administration can change the infrastructure according to its needs. Bringing home seems to depend heavily on externalities.

Both Suchman and Star point out that an increased specialization of work can lead to simplistic representations of work and difficulties with aligning different worlds of work (Star, 1995; Suchman, 1995). They emphasize that a crucial element in the successful use of information technologies is the inclusion and recognition of a broad range of workers and professions who, in one way or another, are affected by the implementation of information technologies.

Much research has stressed the importance of involving and acknowledging the workers and professions whose work are affected by new IT systems and solutions. This was for example a central tenet in the participatory design methods of the 1970's and 1980's (Clement & Besselaar, 1993; Bodker, 1996). However, it is an on-going challenge to align information technologies with the work situation in which they are used. A recent example from Denmark involves the development of a comprehensive electronic health record system (*Sundhedsplatformen*), where several researchers have highlighted the difficulties associated with integrating the many perspectives of different profession including medical secretaries (Bossen et al., 2014; Bossen et al., 2017; Langstrup & Winthereik, 2021).

The three research papers of this dissertation are all concerned with how work becomes represented around the use and development of information technologies. However, the arguments I present are not directly related to the design of information systems, that is they do not necessarily direct towards more or less user involvement in the design phase. Rather, the papers emphasize the organizational relations and priorities which are enacted by information technologies and the surrounding visions

and representations. They address the wider organizational work that has to be acknowledged and taken into consideration in order for information technologies to provide value.

The three research papers ultimately show that the work which is not directly related to the development of data-driven solutions becomes underemphasized and is at risk of being taken for granted and underestimated. When a central vision of the tax administration is to create no-touch customers, how should it handle those taxpayers who do need help from frontline workers? In the article, I show how this decision is passed on to frontline workers, who have to discreetly decide how to handle taxpayers who do need more extensive help. By overemphasizing the value of no-touch customers, the tax administration risks both losing touch with those taxpayers who cannot fulfil the ideal and diminishing the work of frontline workers. The paper, *Organizing Artificial Intelligence,* also highlights how there seems to be an overemphasis and appreciation of the work directly associated with handling data and developing algorithms. While project participants in the risk score project do express concerns about how to include the customs officer's nose in the development of new algorithms, they do not further investigate how the customs officers themselves view their nose. According to the customs officers, their own ability to make thorough inspections has itself become diminished with the introduction of IT-systems and the simultaneous changes made to the organization of their work. This underappreciation of the work not directly related to data-driven solutions is related to technological optimism.

Many researchers have highlighted that the data revolution has been accompanied by an endless amount of hype and hyperbole (Kitchin, 2014; Hockenhull, 2021). Several scholars have warned about believing that the new wave of data technologies will fix the social and political issues of our time (Haraway, 2016, p.3). Evgeny Morozov has used the notion of technological solutionism to characterize the uncritical and unmindful attitude towards recent data technologies, Morozov writes that technological solutionism is:

> Recasting all complex social situations either as neatly defined problems with definite, computable solutions or as transparent and self-evident processes that can be easily optimized – if only the right algorithms are in place! – this quest is likely to have unexpected

> consequences that could eventually cause more damage than the problems they seek to address. (Morozov, 2013, p.5)

Throughout my research, I found that technological optimism was present in the IT projects I followed. In the case of the risk score project, a machine learning algorithm was proposed as a solution to inefficient customs inspection, but it quickly became clear that customs officers were struggling due to decades of downsizing and having little resources available to them. This shows how organizational issues can be cast as technological issues, and exemplifies what McQuillan has called organizations that are data-rich and resource poor (McQuillan, 2018a). Furthermore, as I described in the opening vignette, an explanation algorithm was proposed as a solution to a black boxed machine learning algorithm (which was not yet implemented), which relied on data from a databank (not yet developed), to support customs officers some of whom were not sure they even needed an explanation in the first place. This is a clear example of how technological solutions beget more technological solutions, and can draw attention and resources away from more underlying or pressing organizational issues.

The figure of bringing data home is another example of optimism. This figure suggests that it is possible to separate clearly between public and private actors in the development of data-sharing infrastructures. But as I show through the third research paper, the development of a data-sharing platform inevitably involves reproducing and creating new entanglements between public and private actors. The figure is thus a kind of solutionism, because it casts a complex organizational issue into a simpler form that can easily be 'optimized' or changed. One of the dangers of techno-optimism is that it doesn't recognize that technological solutions depend on existing technologies and the material and organizational conditions in which they are introduced:

> Let us now anathematize techno hubris… It is not enough to put out a new technical infrastructure – it needs to be woven into the daily practices of knowledge workers. It has emerged from the last ten years of information infrastructure development that a wide range of cultural and organizational changes need to be made if the new infrastructure is going to bear fruit (Bowker et al. 2009).

As public organizations continue to place a central emphasis on information technologies in organizational restructurings, it is crucial to question if data-driven

visions and technologies make work situations better or direct resources away from underlying and material conditions central for data-driven technologies to function in the first place.

6.4 Concluding remarks and perspectives for further research

In this dissertation, I have discussed a range of organizational themes and challenges connected to public organizations' pursuit to become data-driven. I have shown that when public organizations work to become data-driven, they are likely to emphasize and focus on specific areas of work, simultaneously relegating other work to the background, with a risk of underemphasizing and ignoring work that is crucial for information technologies and work situations to function. I have also shown that the development of data-driven visions and technologies is involved in redistributing organizational competencies and responsibilities between public employees and departments and between public and private organizations. These reconfigurations continue to be essential for researchers interested in understanding how public organizations are affected by information technologies. STS research and perspectives can exactly be used to highlight the subtle changes and negotiations taking place as public organizations continue to put information technologies central for reimagining their organizations.

At a general level, I believe it is important to keep stressing that while data-driven visions and technologies are ever-present at a discursive level, they are often much more elusive and experimental in organizational settings. There is a large gap between the reports that proclaim that data-driven technologies are pervasive and to the work situations where work is carried out to ensure that public organizations use data to provide organizational value. It continues to be a pertinent challenge for researchers to write about data-driven visions and technologies without reifying them. Following up on this point, I also think we need more studies that nuances and elaborates on the narratives of data-driven and autonomous technologies, which are currently proliferating in public organizations. Future ethnographic studies might benefit from unpacking the broader work situations of data-driven technologies, which are further ahead in their implementation, and presented as stabilized rather than in development.

7 BIBLIOGRAPHY

Aagaard, P. (2017). Den perfekte skandale: Et casestudie af SKAT. In P. Aagaard, & A. Agger (Eds.), *Ledelse I politisk styrede organisationer* (pp. 197-206). Hans Reitzels Forlag.

Accenture. (2019). The Power Of The Data-Driven Enterprise. https://www.accenture.com/_acnmedia/pdf-109/accenture-ao-dde-pov-v5.pdf

Amoore, L. (2009). Algorithmic War: Everyday Geographies of the War on Terror. *Antipode,* 41, 49-69. https://doi.org/10.1111/j.1467-8330.2008.00655.x

Aula, V. (2019). Institutions, infrastructures, and data friction – Reforming secondary use of health data in Finland. *Big Data & Society*. https://doi.org/10.1177/2053951719875980

Balka, E. (2010). Obituary. Susan Leigh Star (1954-2010). *Social Studies of Science*, 40(4), 647–651. https://doi.org/10.1177/0306312710376010

Bang, M. (2006a, August 11) Skat er danmarksmester I digital forvaltning. *Computer world.* https://www.computerworld.dk/art/35129/skat-er-danmarksmester-i-digital-forvaltning

Bang, M. (2006b, August 29) Skat er digitalt forbillede. *Computer world.* https://www.computerworld.dk/art/35399/skat-er-digitalt-forbillede

Bates, J. (2012). "This is what modern deregulation looks like": co-optation and contestation in the shaping of the UK's Open Government Data Initiative. *The Journal of Community Informatics*, 8(2). https://doi.org/10.15353/joci.v8i2.3038

Bates, J. (2014). The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom. *Government Information Quarterly*, 31(3), 388–395. https://doi.org/10.1016/j.giq.2014.02.009

Bates, J., Goodale, P., Lin, Y., & Andrews, P. (2019) Assembling an infrastructure for historic climate data recovery: data friction in practice, *Journal of Documentation*, *75*(4), 791-806. https://doi.org/10.1108/JD-08-2018-0130

Becker, H. (1982). *Art Worlds*. University of California Press.

Bell, D. (1974). *The Coming of Post-Industrial Society*. Harper Colophon Books

Benbouzid, B. (2019). To predict and to manage. Predictive policing in the United States. *Big Data & Society*. https://doi.org/10.1177/2053951719861703

Bigo, D., Isin, E., & Ruppert, E. (2019). *Data Politics: Worlds, Subjects, Rights*. Routledge.

Bilić, P. (2016). Search algorithms, hidden labour and information control. *Big Data & Society*. https://doi.org/10.1177/2053951716652159

Blok, A. & Jensen, T. E. (2009). *Bruno Latour: Hybride tanker i en hybrid verden*. Hans Reitzels Forlag.

Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, N.J: Prentice-Hall.

Bodker, S. (1996). Creating Conditions for Participation: Conflicts and Resources in Systems Development. *Human–Computer Interaction*, 11(3), 215-236. https://doi.org/10.1207/s15327051hci1103_2

Borre, M., Lund, M., & Nielsen, J. (2017, May 16). Kristian Jensen: Kickstartede EFI og fyrede I Skat. *Berlingske*. https://www.berlingske.dk/politik/kristian-jensen-kickstartede-efi-og-fyrede-i-skat

Bossen, C., Jensen, L.G., & Udsen, F.W. (2014). Boundary-Object Trimming: On the Invisibility of Medical Secretaries' Care of Records in Healthcare Infrastructures. *Computer Supported Cooperative Work*, 23, 75–110. https://doi.org/10.1007/s10606-013-9195-5

Bossen, C., N.L.H. Møller, P. Bertelsen, and S. Vikkelsø. (2017, March 17). Lægesekretærers rolle i digitaliseringen overses. *Dagens Medicin*. https://dagensmedicin.dk/369119-2

Bovens, M., & Zouridis, S. (2002). From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. *Public Administration Review*, 62, 174-184. https://doi.org/10.1111/0033-3352.00168

Bowker G.C., Baker K., Millerand F., Ribes D. (2009). Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment. In J. Hunsinger, L. Klastrup & M. Allen (Eds.), *International Handbook of Internet Research*. Springer. https://doi.org/10.1007/978-1-4020-9789-8_5

boyd, D. & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*. 15(5), 662-679. https://doi.org/10.1080/1369118X.2012.678878

Bucher T. (2016) Neither Black Nor Box: Ways of Knowing Algorithms. In S. Kubitschko & A. Kaun (Eds.) *Innovative Methods in Media and Communication Research* (pp. 81-98). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-40700-5_5

Buffat, A. (2015). Street-Level Bureaucracy and E-Government. *Public Management Review*, 17(1), 149-161. https://doi.org/10.1080/14719037.2013.771699

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*. https://doi.org/10.1177/2053951715622512

Callon, M. (1984). Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. *The sociological review*. 32(S1), 196-233. https://doi.org/10.1111/j.1467-954X.1984.tb00113.x

Carse, A., & Kneas, D. (2019). Unbuilt and Unfinished. *Environment and Society*. 10(1), 9-28. https://doi.org/10.3167/ares.2019.100102

Cetina, K. (1995). Laboratory studies: the cultural approach to the study of science. In S. Jasanoff, G. E. Markle, J. C. Peterson & T. Pinch (Eds.), *Handbook of science and technology studies* (pp. 140-166). SAGE Publications. https://www.doi.org/10.4135/9781412990127

Christensen, J. G., & Mortensen, P. B. (2018). *Overmod og afmagt: Historien om Det nye SKAT*. Djøf Forlag.

Clarke, A. E. (2005). *Situational analysis*. SAGE Publications. https://doi.org/10.4135/9781412985833

Clement, A., & Besselaar, P. V. (1993) A retrospective look at PD projects. *Communications of the ACM,* 36(6), 29-37. https://doi.org/10.1145/153571.163264

Creswell, J. W. (2007). *Qualitative Inquiry & Research Design – Choosing Among Five Approaches*. SAGE Publications.

Czarniawska, B. (2004). On Time, Space, and Action Nets. *Organization*, 11(6), 773–791. https://doi.org/10.1177/1350508404047251

Czarniawska, B. (2008). Organizing: how to study it and how to write about it. *Qualitative Research in Organizations and Management*. 3(1), 4-20. https://doi.org/10.1108/17465640810870364

Dalton, C., & Thatcher, J. (2014, May 12). What Does A Critical Data Studies Look Like, And Why Do We Care?. *Society and Space*. https://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care

Danaher, J. (2016). The Threat of Algocracy: Reality, Resistance an Accommodation. *Philosophy & Technology,* 29, 245–268. https://doi.org/10.1007/s13347-015-0211-1

Danish Government (2019) National strategi for kunstig intelligens. https://digst.dk/media/19302/national_strategi_for_kunstig_intelligens_final.pdf

Danish Ministry of Taxation (2004). Det nye skatteministerium. https://www.skm.dk/media/5238/det_nye_skat_juni2004.pdf

Danish Ministry of Taxation (2006). Skat – Fusionen. Det selvangivelsesløse samfund. Afståelsesbegrebet. Aktuelle skattepolitiske initiativer. https://docplayer.dk/1954921-Skat-januar-2006-skatteministeriet.html

Danish Ministry of Taxation (2014). Faldende tillid til SKAT. https://www.skm.dk/aktuelt/publikationer/aktuelle-skattetal/faldende-tillid-til-skat/

Danish Ministry of Taxation (2015a, September 9) SKAT suspenderer al automatisk inddrivelse af gæld [Press release]. https://skat.dk/skat.aspx?oid=2178452

Danish Ministry of Taxation (2015b). Skat ud af krisen – Et oplæg til politisk samling. https://www.skm.dk/media/5034/skat-ud-af-krisen.pdf

Danish Ministry of Taxation (2016a) Et nyt skattevæsen – Ny organisering, flere medarbejdere og velfungerende IT. https://www.skm.dk/media/5032/et-nyt-skattevaesen.pdf

Danish Ministry of Taxation (2016b) SKAT styrker organisationen. Retrieved from: https://skat.dk/skat.aspx?oId=2237572&vId=0

Danish Ministry of Taxation (2017) Fra 1 til 7 styrelser https://www.regeringen.dk/media/3539/fra-en-til-syv-styrelser.pdf

Danish Ministry of Taxation (2020, June 17) Udbytteskandale skal ikke gentage sig. https://www.skm.dk/ministeriet/ministeren/ministerens-debatindlaeg/udbytteskandale-skal-ikke-gentage-sig/

Dourish, P. (2017). *The stuff of bits: An essay on the materialities of information*. MIT Press.

Drucker, P. F. (1969). *The Age of Discontinuity: Guidelines to Our Changing Society*. Butterworth-Heinemann. https://doi.org/10.1016/C2013-0-04383-6

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). *Digital Era Governance: IT Corporations, the State, and e-Government*. Oxford. https://doi.org/10.1093/acprof:oso/9780199296194.001.0001

Edwards, P. N. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. The MIT Press.

Elkjær, B. (2019, November 23). Alle Advarsler og RØDE lamper blev ignoreret I skandalen om udbytteskat. *Information*. https://www.information.dk/indland/2019/11/advarsler-roede-lamper-ignoreret-skandalen-udbytteskat

European Commission (2018). Artificial Intelligence for Europe. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN

European Commission (2020) A European strategy for data. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1582551099377&uri=CELEX:52020DC0066

Flyvbjerg, B. (2017). Introduction: The iron law of megaproject management. In B. Flyvbjerg (Ed.), *The Oxford Handbook of Megaproject Management* (pp.1-18). Oxford University Press. https://ssrn.com/abstract=2742088

Flyverbom, M., & Madsen, A. K. (2015). Sorting Data Out: Unpacking big data value chains and algoritmic knowledge production. In F. Süssenguth (Ed.), *Die Gesellschaft der Daten: Über die digitale Transformation der sozialen Ordnung*. Transcript Verlag.

Forsythe, D. (2001). *Studying those who study us: An anthropologist in the world of artificial intelligence*. Stanford University Press.

Ginsberg, J., Mohebbi, M., Patel, R., Brammer, L., Smolinski, M., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457, 1012–1014. https://doi.org/10.1038/nature07634

Haraway, D. (2016). *Staying with the Trouble: Making Kin in the Chthulucene*. Duke University Press.

Haraway, D. (2018). *Modest_Witness@Second_Millennium. Female Man©_Meets_OncoMouse™* (Second Edition). Routledge.

Hastrup, K., Rubow, C., & Tjørnhøj-Thomsen, T. (2011). *Kulturanalyse: Kort fortalt*. Samfundslitteratur.

Heaphy, L. (2019). Data ratcheting and data-driven organisational change in transport. *Big Data & Society*. https://doi.org/10.1177/2053951719867359

Heide, L. (1994). Punched-card and computer applications in Denmark 1911–1970, *History and Technology*, 11(1), 77-99, https://doi.org/10.1080/07341519408581855

Hockenhull (2020) *Speculative Relations: Data and Digitalization Work in the Danish Public-Private Tech Sector* [Doctoral dissertation, IT University of Copenhagen]. https://en.itu.dk/~/media/en/research/phd-programme/phd-defences/2020/phd-thesis-temporary-version-michael-hockenhull-pdf.pdf?la=en

Hockenhull, M., & Cohn, M. L. (2021). Speculative Data Work & Dashboards: Designing Alternative Data Visions. *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/10.1145/3434173

Holst, H. K. (2016, December 19) Det er kolossalt svigt, at en enkelt medarbejder har siddet med ansvaret alene. *Berlingske*. https://www.berlingske.dk/oekonomi/det-er-et-kolossalt-svigt-at-en-enkelt-medarbejder-har-siddet-med-ansvaret

Høyer, K. (2019). *Hvem skal bruge sundhedsdata–og til hvad?*. Informations Forlag.

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788–8790. https://doi.org/10.1073/pnas.1320040111

Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*. https://doi.org/10.1177/2053951716674238

Irani, L. (2015a, January 15). Justice for 'Data Janitors'. *Public Books*. http://www.publicbooks.org/justice-for-data-janitors/

Irani, L. (2015b). Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk. *South Atlantic Quarterly*, 114(1), 225-234. http://dx.doi.org/10.1215/00382876-2831665

Irani, L. (2015c). The cultural work of microwork. *New Media & Society*, 17(5), 720–739. https://doi.org/10.1177/1461444813511926

Irani, L., & Silberman, M. (2013). Turkopticon. *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 611-620. http://dx.doi.org/10.1145/2470654.2470742

Jensen, C. B. (2010). *Ontologies for Developing Things: Making Health Care Futures Through Technology*. Sense Publishers.

Jessen, C. K. (2018) På 25 år har danskernes syn på skattelettelser ændret sig markant: Ét parti spiller en nøglerolle. *Berlingske*. https://www.berlingske.dk/samfund/paa-25-aar-har-danskernes-syn-paa-skattelettelser-aendret-sig-markant-et

Johansen, H. C. (2007). *Indkomstskatter og offentlig vækst 1903-2005: Dansk Skattehistorie*. Told- og Skattehistorisk Selskab.

Jæger, B. (2020). Digitaliseringen af den offentlige sektor i et historisk perspektiv. In J. Hundebøl, A. S. Pors & L. H. Sørensen (Eds.), *Digitalisering i offentlig forvaltning* (pp. 51-74). Samfundslitteratur.

Kallinikos, J. (2010) Smart Machines. In P. Laplante (Ed.), *Encyclopedia of Software Engineering*.

Keiding (2019) Danskerne vil gerne betale mere i skat for bedre velfærd. *Fagbladet 3F*. https://fagbladet3f.dk/artikel/danskerne-vil-gerne-betale-mere-i-skat-bedre-velfaerd

Kitchin (2014) *The data revolution*. SAGE Publications.

Kitchin, R., & Lauriault, T. (2014). Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work. In J. Eckert, A. Shears & J. Thatcher (Eds.), *Geoweb and Big Data*. University of Nebraska Press (Forthcoming). https://ssrn.com/abstract=2474112

Langstrup, H., & Winthereik B. R. (2021). Infrastrukturstudier. In P. Danholt & C. Gad (Eds.), *Videnskab, teknologi og samfund: En introduktion til STS* (pp. 181-201). Hans Reitzels Forlag.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard university press.

Latour, B. (1996). Aramis, or, the love of technology (C., Porter, Trans.). Harvard University Press.

Latour, B. (2005). *Reassembling the Social - An Introduction to Actor-Network-Theory*. Oxford University Press.

Latour, B. (2013). What's the story? Organizing as a mode of existence. In D. Robichaud & F. Cooren (Eds.), *Organization and organizing: Materiality, agency, and discourse* (pp. 37-51). Routledge. https://doi.org/10.4324/9780203094471

Latour, B., Harman, G., & Erdélyi, P. (2011). *The Prince and the Wolf: Latour and Harman at the LSE*. Zero Books.

Law, J. (1994). *Organizing modernity*. Blackwell.

Law, J. (2004). *After method: Mess in social science research*. Routledge.

Law, J. (2012). Technology and heterogeneous engineering: The case of Portuguese expansion. In W. E. Bijker, P. Hughes & T. Pinch (Eds.) *The social construction of technological systems* (105-129). The MIT Press.

Lazer, D., & Kennedy, R. (2015, January 10). What We Can Learn From the Epic Failure of Google Flu Trends. *Wired*. https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: traps in big data analysis. *Science*, 343(6176), 1203-1205. https://doi.org/10.1126/science.1248506

Löfgren, K., & Webster, C. W. R. (2020). The value of Big Data in government: The case of 'smart cities.' *Big Data & Society*. https://doi.org/10.1177/2053951720912775

Madsen, A. K. (2018). Data in the smart city: How incongruent frames challenge the transition from ideal to practice. *Big Data & Society*. https://doi.org/10.1177/2053951718802321

Maguire, J., & Winthereik, B. R. (2019). Digitalizing the State: Data Centres and the Power of Exchange. *Ethnos*. https://doi.org/10.1080/00141844.2019.1660391

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

McAfee, A., Brynjolfsson, E. (2012, October). Big data: the management revolution. *Harvard business review*. https://hbr.org/2012/10/big-data-the-management-revolution

McQuillan, D. (2016). Algorithmic paranoia and the convivial alternative. *Big Data & Society*. https://doi.org/10.1177/2053951716671340

McQuillan, D. (2018a, September 22). Rethinking AI through the politics of 1968. http://danmcquillan.io/ai_and_1968.html

McQuillan, D. (2018b). Data Science as Machinic Neoplatonism. *Philosophy & Technology,* 31, 253–272. https://doi.org/10.1007/s13347-017-0273-3

Mol, A. (2002) *The Body Multiple: Ontology in Medical Practice*. Duke University Press.

Evgeny Morozov. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. Public Affairs.

Møller, N. H., Bossen, C., Pine, K. H., Nielsen, T. R., & Neff, G. (2020). Who does the work of data?. *Interactions*, 27(3), 52-55.

Narayan, K. (2012). *Alive in the writing: Crafting ethnography in the company of Chekhov*. University of Chicago Press.

National Audit Office of Denmark. (2018). Rigsrevisionens beretning om SKATs kontrol og vejledning på toldområdet. https://www.ft.dk/statsrevisorerne/dokumenter/samling/20171/beretning/sb7/bilag/3/1878268.pdf

Neyland, D. (2008) *Organizational Ethnography*. SAGE Publications.

OECD. (2013). *Tax Administration 2013: Comparative Information on OECD and Other Advanced and Emerging Economies*. http://dx.doi.org/10.1787/9789264200814-en

OECD. (2016a). *Advanced Analytics for Better Tax Administration: Putting Data to Work*. http://dx.doi.org/10.1787/9789264256453-en

OECD. (2016b). *Technologies for Better Tax Administration: A Practical Guide for Revenue Bodies*. http://dx.doi.org/10.1787/9789264256439-en

OECD. (2016c). Rethinking Tax Services: The Changing Role of Tax Service Providers in SME Tax Compliance. http://dx.doi.org/10.1787/9789264256200-en

OECD. (2019). The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies. https://doi.org/10.1787/059814a7-en

Orlikowski W. J., & Scott S.V. (2008). Sociomateriality: Challenging the Separation of Technology, Work and Organization. *The Academy of Management Annals*, 2(1), 433-474. https://doi.org/10.1080/19416520802211644

Overgaard, S. I. (2016). Data Driven Tax Administration [PowerPoint slides]. Danish Customs and Tax Administration. https://www.slideshare.net/siovergaard/data-driven-tax-administration-new-strategy-for-big-data-bi-and-analytics-at-danish-tax-and-customs

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

Pickering, A. (Ed.). (1992). *Science as practice and culture*. University of Chicago Press.

Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2015). *Digital ethnography: Principles and practice*. SAGE Publications.

Plesner, U., Justesen, L., & Glerup, C. (2018). The Transformation of Work in Digitized Public Sector Organizations. *Journal of Organizational Change Management*, 31(5), 1176-1190. https://doi.org/10.1108/JOCM-06-2017-0257

Plesner, U., & Justesen, L. (2020). Managerial Work with Digitalization: A Multi-Sited Ethnographic Approach to Data and Data-Driven Management in Practice. In R. Mir, & A-L. Fayard (Eds.), *The Routledge Companion to Anthropology and Business* (pp. 258-277). Routledge. https://doi.org/10.4324/9781003052456-14

Pors, A. S. (2015). Becoming digital: passages to service in the digitized bureaucracy. *Journal of Organizational Ethnography*, *4*(2), 177 - 192. https://doi.org/10.1108/JOE-08-2014-0031

Prosa (2014) IBM var kildeskattens hemmelige førstehjælper. PROSA. https://www.prosa.dk/artikel/ibm-var-kildeskattens-hemmelige-foedselshjaelper/

Reutter, L., & Spilker, H. S. (2019) The Quest for Workable Data: Building Machine Learning Algorithms from Public Sector Archives. In A. Sudmann (ed.), *The Democratization of Artificial Intelligence: Net Politics in the Era of Learning Algorithms* (pp. 95-107). Transcript. https://doi.org/10.14361/9783839447192-006

Rieder, G., & Simon, J. (2016). Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society*. https://doi.org/10.1177/2053951716649398

Rosenberg, D. (2013). Data before the fact. In L. Gitelman (ed.), *Raw data is an oxymoron* (p. 15-41). The MIT Press.

Ruppert, E. (2012). The Governmental Topologies of Database Devices. *Theory, Culture & Society*, 29(4–5), 116–136. https://doi.org/10.1177/0263276412439428

Schmidt, K. (2016). Reflections on the Visibility and Invisibility of Work. In G. C. Bowker, S. Timmermans, A. E. Clarke & E. Balka (Eds.), *Boundary Objects and Beyond* (pp. 345-351). The MIT press.

Schou, J. (2018). *Remaking Citizenship: Welfare Reform and Public Sector Digitalization.* [Doctoral dissertation, IT University of Copenhagen]. https://pure.itu.dk/portal/files/83711845/PhD_Thesis_Final_Version_Jannick_Schou.pdf

Schou, J., & Pors, A. (2018). Digital by default? A qualitative study of exclusion in digitalised welfare. *Soc Policy Admin*, 53(3), 464-477. https://doi.org/10.1111/spol.12470

Seaver, N. (2014). Knowing Algorithms. In J. Vertesi & D. Ribes (Eds.) *digitalSTS: A Field Guide for Science & Technology Studies* (pp. 412-422). Princeton University Press.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*. https://doi.org/10.1177/2053951717738104

Selinger, E., & Hartzog, W. (2016). Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control. *Research Ethics*, 12(1), 35–43. https://doi.org/10.1177/1747016115579531

Siles, I., Segura-Castillo, A., Solís, R., & Sancho, M. (2020). Folk theories of algorithmic recommendations on Spotify: Enacting data assemblages in the global South. *Big Data & Society*. https://doi.org/10.1177/2053951720923377

Slota, S. C., & Bowker, G. C. (2017). How Infrastructures Matter. In U. Felt, R. Fouché, C. A. Miller & L. Smith-Doerr (Eds.) *The Handbook of Science and Technology Studies* (pp. 529-554). The MIT Press.

Star, S. L. (1985). Scientific Work and Uncertainty. *Social Studies of Science*, 15(3), 391–427. https://doi.org/10.1177/030631285015003001

Star, S. L. (1990). Power, Technology and the Phenomenology of Conventions: On being Allergic to Onions. *The Sociological Review*, 38(1), 26–56. https://doi.org/10.1111/j.1467-954X.1990.tb03347.x

Star, S. L. (1991). The Sociology of the Invisible: The Primacy of Work in the Writings of Anselm Strauss. In D. Maines (Ed.) *Social Organization and Social Process: Essays in Honor of Anselm Strauss* (pp. 265-283). Aldine de Gruyter.

Star, S. L. (1995a). Introduction. In S. L. Star (ed.), *Ecologies of knowledge: Work and politics in science and technology* (pp. 1-38). State University of New York Press.

Star, S. L. (1995b). The politics of formal representations: Wizards, gurus, and organizational complexity. In S. L. Star (ed.), *Ecologies of knowledge: Work and politics in science and technology* (pp. 88-118). State University of New York Press.

Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, 43(3), 377–391. https://doi.org/10.1177/00027649921955326

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology. *Social studies of science*, 19(3), 387-420. http://www.jstor.org/stable/285080

Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information systems research*, 7(1), 111-134. https://doi.org/10.1287/isre.7.1.111

Star, S. L., & Strauss, A. (1999). Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work* 8, 9–30. https://doi.org/10.1023/A:1008651105359

Suchman, L. (1987). *Plans and Situated Actions: The problem of human machine communication*. Cambridge University Press.

Suchman, L. (1995). Making work visible. *Communications of the ACM*, 38(9), 56-64. https://doi.org/10.1145/223248.223263

Suchman, L. (2000). Making a case: 'knowledge' and 'routine' work in document production. In P. Luff, J. Hindmarsh & C. Heath (Eds.), *Workplace studies: Recovering work practice and informing system design* (pp. 29-45). Cambridge University Press. https://doi.org/10.1017/CBO9780511628122.003

Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press.

Suchman, L. (2012) Configuration. In C. Lury & N. Wakeford (Eds.), *Inventive Methods: The happening of the social* (pp. 48-60). Routledge.

Tréguer, F. (2019). Seeing like Big Tech: security assemblages, technology, and the future of state bureaucracy. In D. Bigo, E. Isin & E. Ruppert (Eds.), *Data Politics: Worlds, Subjects, Rights* (pp. 145-164). Routledge.

Tubaro, P., Casilli, A. A., & Coville, M. (2020). The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society*. https://doi.org/10.1177/2053951720919776

Tupasela, A., Snell, K., & Tarkkala, H. (2020). The Nordic data imaginary. *Big Data & Society*. https://doi.org/10.1177/2053951720907107

Vincent Van Roy (2020) *AI Watch - National strategies on Artificial Intelligence: A European perspective in 2019*. European Commission. https://doi.org/10.2760/602843

Veit, D., & Huntgeburth, J. (2014). *Foundations of Digital Government: Leading and Managing in the Digital Era*. Springer. https://doi.org/10.1007/978-3-642-38511-7

Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive science*, 17(1), 7-48. https://doi.org/10.1016/S0364-0213(05)80008-4

Wadmann, S., & Hoeyer, K. (2018). Dangers of the digital fit: Rethinking seamlessness and social sustainability in data-intensive healthcare. *Big Data & Society*, 5(1), 1-13. https://doi.org/10.1177/2053951717752964

Wajcman, J. (2017). Automation: is it really different this time?. *The British Journal of Sociology*, 68, 119-127. https://doi.org/10.1111/1468-4446.12239

Winograd, T. (1993). Categories, disciplines, and social coordination. *Computer Supported Cooperative Work*, 2(3), 191-197. https://doi.org/10.1007/BF00749016

Winthereik, B. R. (2018). Seeing through infrastructure: Ethnographies of HealthIT, Development Aid, Energy and Big Tech. STS Encounters. *DASTS working paper series*, 10(3), 1-26. http://www.dasts.dk/wp-content/uploads/Winthereik-2018-Seeing-through-infrastructure.pdf

Winthereik, B. R. (2020). Is ANT's radical empiricism ethnographic?. In A. Blok, I. Farías & C. Roberts (Eds.), *The Routledge Companion to Actor-Network Theory* (pp. 24-33). Routledge.

Zarsky, T. (2011). Governmental Data Mining and its Alternatives. *Penn State Law Review*, 116(2). https://ssrn.com/abstract=1983326

Ziewitz, M. (2016). Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*, 41(1), 3–16. https://doi.org/10.1177/0162243915608948

Zuboff, S. (1985). Automate/informate: The two faces of intelligent technology. *Organizational dynamics*, 14(2), 5-18. https://doi.org/10.1016/0090-2616(85)90033-6

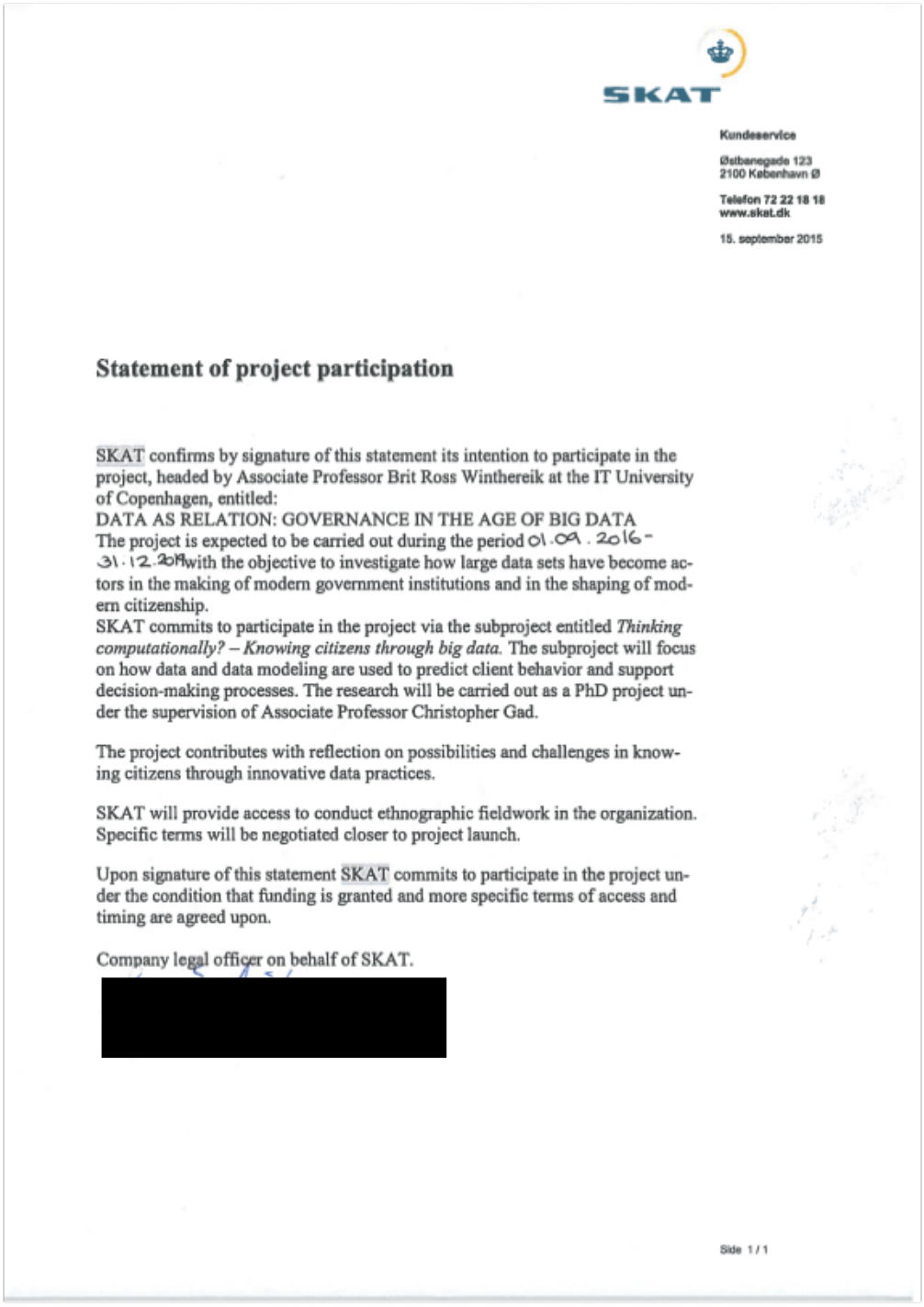
Zuboff, S. (1988). *In the age ofthe smart machine*. Basic Book.

Zuboff, S. (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

Østergaard S.D. (2015) The Danish Tax System and the 'No Touch Strategy'. In C. Gram, P. Rasmussen & S. Østergaard (Eds.) *History of Nordic Computing 4. HiNC 2014. IFIP Advances in Information and Communication Technology* (pp. 58-64). Springer. https://doi.org/10.1007/978-3-319-17145-6_7

8 APPENDIX

8.1 Appendix A. Statement of project participation



SKAT

Kundeservice

Østbanegade 123
2100 København Ø

Telefon 72 22 18 18
www.skat.dk

15. september 2015

Statement of project participation

SKAT confirms by signature of this statement its intention to participate in the project, headed by Associate Professor Brit Ross Winthereik at the IT University of Copenhagen, entitled:
DATA AS RELATION: GOVERNANCE IN THE AGE OF BIG DATA
The project is expected to be carried out during the period 01.09.2016-31.12.2019with the objective to investigate how large data sets have become actors in the making of modern government institutions and in the shaping of modern citizenship.
SKAT commits to participate in the project via the subproject entitled *Thinking computationally? – Knowing citizens through big data.* The subproject will focus on how data and data modeling are used to predict client behavior and support decision-making processes. The research will be carried out as a PhD project under the supervision of Associate Professor Christopher Gad.

The project contributes with reflection on possibilities and challenges in knowing citizens through innovative data practices.

SKAT will provide access to conduct ethnographic fieldwork in the organization. Specific terms will be negotiated closer to project launch.

Upon signature of this statement SKAT commits to participate in the project under the condition that funding is granted and more specific terms of access and timing are agreed upon.

Company legal officer on behalf of SKAT.

[redacted]

Side 1 / 1

8.2 Appendix B. Examples of interview guides

Interview Guide - Kundeservice SKAT

Tidspunkt: Mandag d. 29 august
Interviewer: Bastian Jørgensen
Informant: [redacted] (telefonagent og intern coach i forhold til kundebetjening)

Projekt beskrivelse

Hvordan har digitaliseringen påvirket forholdet mellem SKAT og de danske borgere?

Emne for interview: Digitalisering og kommunikation over telefon.

Spørgsmål

1. Hvordan har du oplevet digitaliseringen i forhold til dit arbejde i SKAT?

2. Har digitaliseringen ændret forholdet mellem SKAT og borgerne?

3. Hvad er dit indtryk af borgernes generelle holdning til SKAT?

4. Hvilke hjælpemidler og "signaler" bruger du i kommunikationen med borgerne (logs,borgerens tone i telefonen, årsopgørelsen)?

1

Interview:
Længde: 60 minutter
Interviewform: Semistruktureret interview
Valg af interviewpersoner er foretaget på baggrund af ønsket om 1) at snakke med de mest involverede projektdeltagere i GUL og 2) et ønske om at få perspektiver fra deltagere med forskellige organisatoriske baggrunde.

Baggrund:
Formålet med denne undersøgelse er at belyse og beskrive organisatoriske forandringer i SKAT i forbindelse med organisationens øgede fokus på IT og data. Undersøgelsen udgangspunkt er et etnografisk studie af IT projektet GUL (generel udstillingsløsning), hvis formål er at udvikle en udstillingsplatform, hvorfra SKAT kan dele data til eksterne kunder.
I projektet GUL er der en stor fokus på at arbejde med agile IT udviklingsmetoder som f.eks. Prince2Agile, SCRUM og DevOps, hvilket står i kontrast til tidligere udviklingsparadigmer i det offentlige. Derudover ønskes der at netop udviklingen og driften af IT løsninger i højere grad skal kunne klares af medarbejdere fra SKAT. Dette er blandt andet et forsøg på at gøre op med tidligere afhængigheder til eksterne virksomheder og konsulenter. GUL er interessant, da det på mange områder er et foregangsprojekt, der forsøger at kombinere nye idéer om IT udvikling og skabe samarbejde mellem nye IT afdelinger i SKAT men også til andre offentlige og private organisationer. Igennem observationer og interviews med projektdeltagere i GUL ønsker denne undersøgelse at diskutere, hvilke forandringer og idéer der er i en offentlig organisation, hvordan disse forandringer forsøges at implementeres i praksis og hvilke organisatoriske udfordringer og forhindringer der eksisterer.

Interviewguide (Standard til SKAT medarbejdere)

1. **Din baggrund i SKAT & GUL**
2. **Organisatoriske forandringer i forbindelse med GUL**
   - 2.1. Forandringer i forbindelse med nye udviklingsmetoder
   - 2.2. Forandringer i forbindelse med deling af data
   - 2.3. Nye samarbejder (internt og eksternt)
3. **Fremtiden med GUL**
4. **Om et øget fokus på IT og data i SKAT (hjemtagelse af data, det datadrevne skattevæsen)**
5. **Evt.**

Interviewguide (Liste med spørgsmål)

1. **Din baggrund i SKAT & GUL**
    a. Hvor lang tid har du været ansat i SKAT?
    b. Hvilken afdeling og stilling er du ansat i?
    c. Hvad er din rolle eller tilknytning til GUL projektet?

2. **Organisatoriske forandringer i forbindelse med GUL**

    a. Er GUL et typisk projekt i SKAT? Hvorfor/Hvorfor ikke?
    b. Hvad er det mest interessante ved GUL for dig?
    c. Hvad skal GUL gøre SKAT i stand til?
    d. Hvilke værdier for offentlig administration repræsenterer GUL?

2.1 Forandringer i forbindelse med nye udviklingsmetoder

    e. Hvad betyder det for dig at arbejde agilt?
    f. Er GUL ifølge dig et agilt projekt? Hvorfor eller hvorfor ikke?
    g. Hvad betyder det for dig at arbejde med DevOps?
    h. Hvordan ser du sammenspillet mellem agile metoder, SCRUM og DevOps?
    i. Hvordan passer agile metoder, SCRUM og DevOps sammen med statens IT projektmodel?
    j. Er der specifikke udfordringer med at arbejde agilt i en offentlig organisation?
    k. Hvordan vælger man imellem mange forskellige software muligheder? (tænk DevOps toolchain, tænk API management platform.)

2.2 Forandringer i forbindelse med deling af data

    l. Hvordan forandrer GUL måden hvorpå SKAT deler data med eksterne parter?
    m. Hvorfor ønsker man at ændre måden hvorpå SKAT deler deres data?
    n. Hvad er de største udfordringer i arbejdet med at kunne dele data direkte fra SKAT?
    o. Hvorfor ønsker SKAT at hjemtage data? Hvilke idéer ligger bag dette ønske?
    p. Hvad er udfordringerne med at ”hjemtage” data (f.eks. fra KMD og DXC)?

PART II

RESEARCH PAPERS

Research paper 1:

Helping or intervening? Modes of ordering in public sector digitalization

Jørgensen, B., & Schou, J. (2020). Helping or intervening? Modes of ordering in public sector digitalization. *Journal of Organizational Ethnography* (pp. 265-279), 9(3). https://doi.org/10.1108/JOE-02-2019-0015

Note: This paper is co-authored with Jannick Schou, and it includes empirical examples from citizens service centers (Jannick's empirical site), as well as empirical examples from call centers in the DCTA.

Helping or intervening? Modes of ordering in public sector digitalization

Bastian Jørgensen and Jannick Schou
Department of Business IT, IT University of Copenhagen, Copenhagen, Denmark

Received 19 February 2019
Revised 17 August 2019
1 November 2019
31 January 2020
Accepted 2 February 2020

Abstract

Purpose – This paper examines how digital reforms affect the relationship between frontline workers and citizens in Danish public sector institutions. Using ethnographic research in two branches of public administration, the study highlights how frontline workers act in accordance with seemingly contradictory modes of ordering. Their acts problematize linear conceptualizations of change that often prevail in digital reforms.

Design/methodology/approach – The paper is based on a comparative ethnographic study of frontline workers in the Danish tax and customs administration and municipal citizen service centers. The concept of modes of ordering is used to highlight new tensions that arise as frontline workers adapt to make digital reforms work.

Findings – Frontline workers act according to two different modes of ordering based on the separation between helping citizens help themselves and helping citizens directly. National policies and strategies promote the underlying rationale of the first mode but, as this paper shows, this mode is sustained by a second mode, which involves the intervention of professionals when citizens cannot be helped to help themselves.

Originality/value – The paper, which contributes to our understanding of how digitalization is changing public administrations and the relationship between frontline workers and citizens, challenges applying a linear, technocratic focus in discourses on public sector digitalization and highlights the contradictory practices of frontline work. It demonstrates the necessity of going beyond policy narratives and calls for increased attention to how frontline workers adapt to make reforms work.

Keywords Frontline work, Digitalization, Modes of ordering, Discretion, Public sector reform, Ethnography

Paper type Research paper

Introduction

"Normally, I think it's really, really rewarding," explains Maria enthusiastically[1]. She has been employed as a frontline worker at a Danish citizen service center (CSC) for several years. Located in a mid-sized municipality in Zealand, Denmark, the center is only one of many established in Denmark's 98 municipalities since the middle of the 2000s "I think it's fun to try and figure out their way through the system," she continues, referring to what takes up a large part of her everyday work: helping people use government websites to request welfare benefits and other public subsidies. Since the mid-2000s, successive Danish governments have advanced the use of digital self-service solutions across most areas of welfare provision (Schou and Hjelholt, 2018). People are expected to help themselves online, with frontline workers in CSCs expected to help those who cannot help themselves (Pors, 2015). Framed as a way of improving quality, enhancing efficiency and making people self-sufficient, public sector reforms have made digital platforms "mandatory for more than 100 administrative procedures" (The Government, Local Government Denmark and Danish Regions, 2016, p. 6).

"But I often get frustrated on their behalf," adds Maria, adjusting her tone somewhat and adding, "Many of the people we talk to expect us to be understanding. If you know that someone has struggled with a request for 45 min, you've tried to help them several times, and

This work was supported by Velux Fonden [Grant: 12823 (Data as Relation: Governance in the Age of Big Data)].

Journal of Organizational Ethnography

2046-6749
DOI 10.1108/JOE-02-2019-0015

then they press the last button only for the system to respond that there's been a technical error. Well, then you just feel exhausted." Earlier in the same conversation, Maria recounted how the official systems often do not work as intended. People get mired in malfunctioning websites or bureaucratic procedures that endlessly toss them around the administrative system. "So most of the time, it's fun, and at times it's frustrating," she concludes, "especially on behalf of citizens. Because you've experienced the same situation yourself. You've been in their shoes. So, in many ways, we're acting on people's behalf."

This brief narrative might seem fairly uneventful or even mundane. We learn about Maria, a frontline worker trying to help citizens to the best of her ability. She often shares their frustration whenever something does not go according to plan. We learn how she sometimes acts on behalf of citizens in these situations. Despite the apparent ordinariness of the encounter, it is quite surprising in terms of recent political strategies and reforms as frontline workers in CSCs (but also in many other branches of the Danish public sector) acting on behalf of citizens has become increasingly rare. Since the early 2000s, policymakers have pushed for the exact opposite behavior. Through wide-ranging digitalization reforms, citizens have increasingly been expected to communicate with the government through digital interfaces due to the fact that self-service platforms have generally replaced face-to-face encounters.

How, then, are we to understand the episode recounted above? And, more generally, how does the relationship between frontline workers and citizens change and develop as public administration becomes increasingly more digital? This paper addresses these questions by taking an ethnographic approach to examining encounters between frontline workers and citizens in the Danish public sector. Exploring how welfare work takes place at two distinct sites – CSCs and Danish tax and customs administration (DTCA) service centers – the paper showcases some of the tensions and dilemmas that arise in the professional work of frontline workers as they navigate issues that involve helping people in their everyday work. Using the concept of *modes of ordering* (Law, 1994) as a foundation for understanding the often conflicting and tension-filled conditions frontline workers have to negotiate on a day-to-day basis, this study provides empirical insights into how the relationship between frontline workers and citizens is modified and transformed at the forefront of digitalization by combining ethnographic findings from the two sites. Our study adds to contemporary ethnographic accounts of public reforms and administration (Boll and Rhodes, 2015; Fassin, 2015; Bjerge and Row, 2017). Combining work on street-level bureaucracy (Maynard-Moody and Musheno, 2000; Lipsky, 2010; Evans, 2016; Hupe *et al.*, 2016) with science and technology studies (Law, 1994), we explore the idea that digitalization automatically implies a smooth, epochal shift from street-level bureaucracy to system-level bureaucracy (Bovens and Zouridis, 2002). Consequently, we focus on the oft-neglected work that goes into making bureaucracies digital, showcasing how frontline workers do not merely follow policy ideals but engage in creative workarounds to help people who would have otherwise be left on their own. This study demonstrates that, rather than being a hindrance to the supposedly inevitable progress of technological change, the movement between different modes of ordering makes it possible for increasingly digitalized institutions to function at all.

Vision of the future: the disappearance of frontline workers?

Since the early 1990s, public administrations have turned to digitalization as a way of improving public services (Chadwick and May, 2003; Dunleavy *et al.*, 2006; Margetts, 2009; Pollitt, 2011). Information and communication technologies (ICTs) are seen as necessary to keep up with the demands imposed both by citizens and societal developments. Policymakers in Denmark have also been keen on using ICTs (particularly self-service platforms and websites) to simultaneously cut costs and to improve the efficiency of welfare institutions and the quality of service delivery (Jæger and Löfgren, 2010; Jæger and Pors, 2017; Pors, 2015).

Since the early 2000s, Denmark has pursued a strategy based on citizens being digital by default, replacing previous interpersonal encounters with digital platforms. Today, the country is often ranked as having one of the most digital administrations in Europe, with official benchmarking exercises promoting the country as an international leader in shaping a comprehensive digital agenda (EC, 2019).

A variety of umbrella terms have been used in the study of these types of changes, including: e-government, e-governance, and digital era governance, which are just a few of the ways digital forms of public administration are increasingly referred to as (Dunleavy *et al.*, 2006; Henman, 2010; West, 2005). Existing studies often emphasize the centrality of ICTs in changing governments and welfare practices. As Dunleavy *et al.* (2006) write: "The advent of the digital era is now the most general, pervasive and structurally distinctive influence on how governance arrangements are changing in advanced industrial states" (Dunleavy *et al.*, 2006, p. 225). Similar ideas are often found in official policy documents and reports (author: year), framing digital technologies as increasingly more fundamental to state institutions and governance.

These portrayals of managerial shifts have often been underpinned by what we call a linear orientation towards the future, rather than a focus on how ICTs actually impact governments. Indeed, when it comes to the relationship between government and its citizens, the focus is often on the implementation of zero-touch technologies rather than the changing practices of frontline workers and their everyday practices. This conceptualization of linear change involves a vision of the gradual disappearance of frontline workers and the simultaneous emergence of citizens with the ability to handle their affairs and issues with the government – without direct contact to a frontline worker. As discussed later in this article, these ideas have also been prevalent in the policies and strategies that have shaped Danish CSCs and DTCA service centers. The vision involves a more or less epochal shift from a supposedly bureaucratic, slow, and inefficient past to a fast, efficient, and smooth future, buttressed in every respect by digital platforms.

In contrast to this high-level focus on shifts in management ideas, studies inspired by Michael Lipsky's (2010) work on street-level bureaucracy have turned to the ways in which government use of ICTs has an impact on the relationship between frontline workers and citizens (Bovens and Zouridis, 2002; Jansson and Erlingsson, 2014; Buffat, 2015; Pors, 2015; Hansen *et al.*, 2018) These studies draw specific attention to the *discretion* of frontline workers, which is understood in this study as the ability of frontline workers to act according to their best judgment rather than solely following established guidelines. For Lipsky, the ability of street-level bureaucrats to act with discretion is a defining element of their work. His emphasis on this concept radically challenged the view of street-level bureaucrats as simply implementing policies from above, helping them to instead be portrayed as policymakers in their own right (Lipsky, 2010; Evans, 2016). As Maynard–Moody and Musheno have demonstrated, street-level bureaucrats often occupy a paradoxical professional position in which they end up acting both as agents of the state and agents of citizens (Maynard-Moody and Musheno, 2000; see also Fassin, 2015; Zacka, 2017).

While existing research on how digital reforms affect street-level bureaucracy has offered important insights into the workings of everyday administrative work, many studies have nonetheless tended to presuppose a linear conceptualization of change similar to that of those studying larger managerial changes. Bovens and Zouridis (2002), for example, argue that ICTs enable a more or less wholesale transition from first street-level bureaucrats to screen-level bureaucrats and, subsequently, system-level bureaucrats. That is, according to these authors, a transition from bureaucrats who meet citizens directly and have substantial discretionary power (street-level) to bureaucrats who interact with citizens through information technologies and whose role it is to fill out electronic forms (screen-level) followed by bureaucrats who develop, maintain, and optimize information systems where

transactions "have been fully automated" (2002, p. 179). In this imagined scenario, direct communication between bureaucrats and citizens is no longer needed, as interfaces replace former points of contact. The idea that the discretion of frontline workers disappears as ICTs become part of the interaction of frontline workers with citizens has been labeled and discussed as the curtailment thesis (Buffat, 2015).

To our mind, there is a need to move beyond these somewhat static ideas of discretion as simply vanishing as ICTs become evermore prevalent. Instead, our ethnographic work has prompted us to describe how the relationship between frontline workers and citizens has become reconfigured and modified as ICTs are introduced into already existing bureaucratic situations. Rather than a linear change from one form of bureaucracy to another, our study highlights how new tensions and professional dilemmas arise and are negotiated on a day-to-day basis. We show how multiple modes of ordering are at play at the same time, sustaining, undermining, and complementing each other. In doing so, this study provides an alternative understanding to conceptualizations of linear technological change and the gradual disappearance and/or diminishing role of frontline workers that existing research often underlines. At the same time, our study adds new dimensions to the tensions explored by the literature on street-level bureaucracy (Maynard-Moody and Musheno, 2000; Zacka, 2017) by specifically showcasing how policy ideals become sustained through seemingly subversive practices and contrasting modes of ordering. Rather than moving from one form of bureaucracy to another, we instead identify the emergence of new complex and layered practices.

Fieldwork

This paper takes an ethnographic approach to the study of administrative practices in accordance with recent methodological interventions by political science and public administration scholars (Dubois, 2009; Boll, 2015; Boswell and Corbett, 2015; Herzog and Zacka, 2017; Zacka, 2017). Combining observations with document analysis and interviews, ethnographic research allows us to examine the relationship between official policy narratives and the work these policies aim to define and change. To avoid the trap of being too anecdotal and myopic, we believe that ethnographic observations provide important insights into how political reforms are being made to work (see also Brodkin and Marston, 2013; Rhodes, 2014; Brodkin, 2017). For this study, we draw on ethnographic research on digitalization carried out in two public sector institutions: DTCA and CSCs.

The DTCA fieldwork comprised approximately 80 days of data collection carried out by the first author between 2016 and 2018, its overarching interest the tax administration's recent vision to become a data-driven organization. In 2016, the first author observed how frontline workers communicated with taxpayers on the phone, using social media and in e-mails. This article draws particularly on data and fieldnotes from observations conducted at a DTCA service center, where the first author listened to telephone conversations between frontline workers and the Danish taxpayers they were assisting. This involved sitting next to the frontline worker and listening to the conversation on a headset, making it possible to hear both speakers. The phone calls were often followed by a short break, providing the opportunity to discuss what was said with the frontline worker. Sitting next to the frontline worker also made it possible to observe which caller documents on the computer were used and when as the conversation took place. In addition, the article draws on seven interviews with managers, frontline workers, and customer strategists working directly with how digitalization has affected the relationship between DTCA and taxpayers.

The second author conducted ethnographic studies at CSCs from early 2017 to early 2019, collecting material that includes a combination of interviews with frontline workers and managers in nine different CSCs (seven visited in 2017 and two in 2019) across the five Danish regions, as well as ethnographic observations conducted in an additional CSC for five months

from late 2018 to early 2019. The interviews, which made it possible to gain a sense of the everyday practices taking place across the various regions and municipalities, focused on the daily experiences frontline workers had interacting with citizens. The observations, meanwhile, were conducted after the interviews in a single location and consisted of observations, informal conversations, and semi-structured interviews. The second author followed the day-to-day practices of frontline workers by shadowing them as they engaged with and helped people. After each encounter, the author was able to ask follow-up questions and talk with the frontline worker about what occurred.

In this study, we combine ethnographic material from these two sites to provide insights into how digital reforms are creating new tensions in professional work and how multiple modes of ordering are negotiated, performed, and practiced on a daily basis. Digitalization reforms implemented in the Danish public sector since the early 2000s have affected both sites. At both sites, interactions with citizens have changed, as contact has moved online, and frontline workers are expected to help citizens use digital platforms. Even so, DTCA and CSCs might appear to differ at first sight in that DTCA service centers are nationally regulated and function as call centers, while CSCs are municipal-level institutions that primarily deal with walk-ins and face-to-face encounters. Nonetheless, our research shows that not only are the practices remarkably similar across the two sites, but the workarounds and tensions we identified are also similar. Using these two sites in tandem thus provides an opening for understanding some of the inherent tensions in large-scale public sector digitalization efforts, allowing us to zoom in on the both conflicting and complementary modes of ordering underpinning the transition to increasingly digitalized welfare encounters.

Analytical approach

In our analysis, to emphasize the tensions and contradictory ways public sector reforms are implemented, we draw on Law's (1994) concept of modes of ordering, introduced in his 1994 book, *Organizing Modernity*. He suggested that attempts at ordering are never complete and create new and unintended forms of order (Law, 1994, p. 2). According to Law, modes of ordering "are recurring patterns embodied within, witnessed by, generated in and reproduced as part of the ordering of human and non-human relations" (1994, p. 83). Modes of ordering allow actors in the field to make sense of and account for their actions, as well as the actions of colleagues. But they also prescribe how actors could and should act (for an application of modes of ordering in social work, see Elgaard, 2017). In this sense, modes of ordering allow us to move beyond dual notions of agency and structure. Indeed, we use the idea of modes of ordering – and its underlying relational approach – as a way of analyzing patterns that resonated across the ethnographic work we conducted among managers, frontline workers, citizens, and institutional spaces. In analyzing our material, we discuss two modes of ordering that emerged in our analysis, namely: help to self-help and directly helping. These modes linked together a series of observations that, at first sight, appeared contradictory and incommensurable; yet, as we will demonstrate, they nonetheless functioned in highly complementary ways.

Using the notion of modes of ordering allows us to analyze how professional judgment and discretion play out in light of public institutions being affected by digital reforms by shifting the analytical focus slightly. Frontline workers exerting discretion and professional judgment presupposes that certain policies or rationales for action are in place. Maynard and Moody (2000) argue that discretion is inevitable as, "[r]ules and procedures can never universally fit each individual and every circumstance, so judgements must be made" (p. 338). Yet, whereas discretion primarily brings attention to the actions of frontline workers, modes of ordering direct our attention to the interplay between individual actions, political contexts, institutional constraints, citizens and the various artifacts that are part of and underpin

particular modes. For DTCA frontline workers, for example, it is technologically possible to edit the taxpayer documents but, ideally, they should educate taxpayers to do it themselves. In CSCs, there is nothing that constrains frontline workers from taking over someone's mouse or keyboard. And so, at times, they do just that. Understanding how discretion and professional judgment play out in such situations requires taking these seemingly mundane practices into account. Thus, we suggest that not only do modes of ordering provide us with a productive heuristic for ethnographic research, but they also connect to existing work on street-level bureaucracy. What is often described as different forms of discretion, coping mechanisms, and creative strategies (Lipsky, 2010; Tummers *et al.*, 2015) can be seen as part of (and in response to) particular modes of ordering. This allows us to see the deployment of discretion, coping mechanisms, and creative strategies as more than a matter of individual abilities. To understand these, we must place them in their relational context, as part of wider emergent, recurrent, and contingent patterns of interaction stabilized over time. With modes of ordering, then, our aim is to slightly shift the attention and emphasize how such responses are made possible and emerge in the interplay between frontline workers, citizens, technologies, and new organizational formations.

Case 1: DTCA service centers

> July 8, 2016. I'm set to visit one of the DTCA service centers. Even though I was told in advance that the center was located just next to the train station and that I could not miss it, I have spent forty minutes looking for the entrance. There do not seem to be any signs anyway nearby pointing to a tax center, nor are any of the adjacent buildings marked with the DTCA logo, a yellow circle with a royal crown. Eventually, however, I do manage to find it. The entrance is hidden in a basement parking garage under a large concrete and glass building. As I enter the building, I'm struck by how busy everyone in the building seems to be. I realize that, contrary to my own inability to locate the center, taxpayers are already here [though not physically]. They are being helped and guided by tax employees on the phones and with computers. (Fieldnotes, 2016, first author)

Nowadays, DTCA service centers like the one just described represent the main entryway to DTCA. At these centers, frontline workers are not available to speak face-to-face with taxpayers. Instead, they communicate with taxpayers from a distance via telephones and computers to help them fill out their official tax returns and forms.

Centers like these are relatively new, and tax administration managers often connected their emergence to either digitalization or the structural reforms undertaken in the Danish public sector in recent decades. A manager in the center who had been employed in the tax administration since the 1980s said digitalization was a rather difficult notion to pinpoint and many of the changes in communication between taxpayers and frontline workers were a consequence of a larger structural reform implemented in 2005. Prior to the reform, local municipalities were responsible for guiding taxpayers with reporting their taxes. This service often relied on a close relationship between frontline workers and taxpayers. In small municipalities, the manager explained, they often got to know each other. In a somewhat nostalgic tone, he recounted the time when communication between taxpayers and frontline workers took place in the municipalities. He remembered a frontline worker who had been so well-liked that some people asked specifically for her help every time. If she was not at work that day, they preferred to come back another time.

In 2005, these kinds of relationships changed, the manager said, as the structural reform took hold. Municipal tax administrations were merged with national customs and tax authorities into one central administration called SKAT (Danish Ministry of Taxation, 2006, p. 6). A centralized tax administration was justified as being more efficient and able to provide better services. At that time, keeping local customer services in the municipalities was emphasized as important because they were supposed to continue giving taxpayers guidance

(Danish Ministry of Taxation, 2006., p. 7). Over time, however, the idea that local support functions should be maintained has gradually disappeared, with face-to-face administrative assistance becoming increasingly rare. Visiting one of the service centers, these changes were highly visible:

> After listening in on a telephone call between a frontline worker and a taxpayer, the frontline worker turns to me and asks: "Well was not she kind? She's a good example of how not everyone is satisfied with digitalization. She cannot do everything herself now that she cannot go down to a tax office and talk to someone." I ask if the taxpayer's problem had anything specific to do with digitalization and whether it had not always been difficult to report your taxes. The frontline worker is quiet for a moment, as if to gather her thoughts, and then replies: "It's many things . . . Earlier, the elderly had the opportunity to go down to their local tax centers . . . At the tax centers, someone helped them and you often had a personal relationship. That meant a lot, but it's not possible anymore. Now almost all tax centers are closed, and it is only possible to make an appointment at very few of them. Our instructions are that if we can handle people on the phone, then we should do that. We should not overburden the tax centers. If we did, it would not have made any sense to close them in the first place." (Fieldnotes, 2016, first author)

This conversation begins to draw the contours of the ideal image of taxpayers, and with it, a specific mode of ordering. The ideal taxpayer is someone who does not need to communicate directly with a frontline worker but is self-sufficient and capable of handling their tax affairs alone – what tax administration employees also refer to as no-touch customers. Based on this ideal taxpayer, the tax administration has become organized around a particular mode of ordering that emphasizes self-service platforms, telephones, e-mail, and social media as the main means of interaction, and frontline workers are instructed to deal with people who are not self-sufficient without turning to face-to-face encounters. A statement by an employee who worked with customer strategies emphasizes this particular ordering of relations:

> Within groups where it is appropriate and realistic, we want people to search for information in order to help themselves. But it can escalate and then we'll manage them on the telephone, with e-mail and on social media and, in the worst case, we can arrange a physical meeting, which is the most expensive and difficult to manage for us. But it can be the right thing to do with some customers. . . We have a democratic responsibility, and it can be the right thing to help people by choosing a physical meeting. We just have to limit the service, so we do not waste time on expensive tools for people who do not need it.

The strategist highlights how the ideal taxpayer is someone who does not need to be in direct contact with a frontline worker. When taxpayers seek further help or guidance, the strategist describes the situation as escalating. This resonates with ideas promoted by large political organizations such as the OECD. In a 2014 report, *Increasing the Use of Self-service Channels*, the OECD argued that promoting self-sufficient taxpayers is essential to reduce the cost of tax administrations. Self-service channels are described as an alternative to "traditional and more expensive forms of taxpayers services" (OECD, 2014, p. 3). The mode of ordering – and the notion of an ideal taxpayer that it contains – appears to rely on a similar image.

Yet, from the account provided by the strategist, we also learn that the tax administration has to take other principles into account, namely those linked to democratic responsibility. While the tax administration does rely on a certain mode of ordering and image of the taxpayer, it must simultaneously deliver services that do not fit these ideal ways of ordering relationships. This is a paradoxical situation in which an ideal is created for taxpayers, who are imagined as being able to help themselves, while at the same time, the administration creates mechanisms for dealing with taxpayers who do not fit this ideal. The ideal taxpayer is both a universal and particular figure at the same time.

Ideal or not?

Moving back to the encounter with the frontline worker narrated earlier, we learn how these tensions and ideals are navigated in daily practice. Drawing on her professional judgment, it is up to the frontline worker to solve the tension between the ideal taxpayer and the practicalities of helping people with their problems, here and now. She is instructed to help as many people as possible on the phone to avoid overburdening the DTCA service centers. Yet, while the situation is in many ways highly ordered, the frontline worker still needs to navigate and find ways around the official ideal. She needs to carve out a space for handling the situation in other ways than those formally described by policymakers and strategists:

> The tax administration's attitude is that if you're digital, you should be able to handle everything yourself. But we are, after, all still a service center. So you have somebody who you can call if things are too difficult. People are always really thankful when we have helped them. That's the pleasure of working in a call center, when you can feel that you've helped people and they're happy.

Thus, we can thus begin to see a tension between quite two different modes of ordering – two ways of relating to and ordering citizens. The first revolves around the idea that "if you're digital you should be able to handle everything yourself" (as stated by frontline worker). This materializes in the form of concrete instructions as to how frontline workers should assist taxpayers by providing guidance. It is actualized through the geographical distance between frontline workers and citizens, made possible by using phones, social media and other communication platforms. This is a mode of ordering that is premised on the expectation that taxpayers are capable of helping themselves. Following this mode, frontline workers should gently guide taxpayers and assist them in finding the right path on their own. But they should not make the changes *for* them. However, it is not always easy for frontline workers to refrain from solving the problem for taxpayers. "I think that we sometimes do a bit more than we should," a frontline worker noted, "and that maybe we could have guided them, but in some cases, we just know that it's going to take a long time for them. Also, if we do not guide them all the way to the end, they might make a mistake and make some changes that turn out to be fatal for them. We're not interested in that either." Thus, while the official instructions are to *guide* taxpayers – that is, *not* fill out the tax documents for them – frontline workers do not always and cannot always act accordingly. As the frontline worker notes, she is not always acting according to the official instructions. She sometimes does more than she should.

Filling out tax papers for taxpayers is a way of solving some of the issues that emerge whenever taxpayers encounter too many difficulties. An additional way of relating to people thus starts to take over whenever the first way breaks down or seems impossible to realize. In this second mode, what is at stake is not an attempt to turn citizens into an ideal. Instead, it is a more direct form of intervention that seeks to solve the administrative problem here and now. This act of *doing it for them* would seem to undermine the idea that all taxpayers should learn to become digital. Yet, this subtle shift in practice could also be seen as the key to what actually makes the system function. It allows some taxpayers to become digital, even though they cannot manage to use the system according to the ideal of a taxpayer, as a frontline worker explains:

> I do not know if it could be done otherwise, but it's because, in this digital world, you want everything to be digital. Sometimes we're stricter with young people, where we say: "If you just log on, then I will show you how you do it and then we can figure it out". But there are also tax cards and stuff like that, where I feel that if the taxpayers fumble around and make a mistake then I'll have a bad conscience because I did not help them, and then I do it for them. Sometimes you guide them all the way through, and then a year or two passes by, and they've forgotten everything about how to do it anyway.

The frontline worker thus seems to be caught in a double bind: the worker feels guilty if the taxpayer fumbles around and will often decide to help the taxpayer directly to mitigate such a situation. Yet, in the very act of helping, a new form of guilt seems to arise, as this is not how the worker is expected to act.

Case 2: CSCs

CSCs have been reorganized along similar lines as the tax administration. Influenced by many of the same reforms as call centers, CSCs have also seen a pronounced shift in how frontline workers are expected to help people (on this, see Pors, 2015). Bhatti *et al.* (2011) recount how CSCs were seen as an integral part of "putting citizens at the center," one of the major slogans driving the structural reform that took place in Denmark in 2007. At the time, "services undertaken in the CSCs [were], to a large extent, social services" (p. 5) in which the "citizen can show up in person to access multiple municipal services (including social services)" (Bhatti *et al.* (2011)). They argue that "[t]he services the centers provide concentrate on the sort of things that many citizens need. [. . .] [T]he CSCs are one-stop shops where the services most citizens need are concentrated in one place" (p. 5). CSCs have transformed significantly since their inception (Pors, 2015). Digitalization reforms have led to almost all administrative areas previously handled in CSCs being moved online. Instead of showing up physically at centers, people are expected to be able to handle administrative tasks using digital platforms.

As one frontline worker explains in a 2019 interview:

> It has changed a lot. [. . .] While they [the citizens] used to be serviced very, very much, this has been phased out in favor of ... Well, of course, we service them as well as we can ... But it has changed because people need to do it themselves. Today, people need to do it themselves. At that point [eight years ago], all information passed through the information desk, [and] from Monday to Thursday, people used to stand in long lines all the way down the street. They do not do that anymore. (Interview with frontline worker, 2019)

Both frontline workers and managers often stated that citizens have increasingly been expected to manage tasks previously handled by frontline staff through digital self-service platforms. This has impacted how citizens and frontline workers encounter one another. Previously, citizens were expected to use CSCs, often standing in long lines, but now only people who cannot do it alone online show up. Yet, as most CSCs have been downsized with the advancement of digital platforms, there are still lines – albeit comprised of a very different group of citizens.

"You take a number," a frontline worker explained, recounting how these encounters now take place, "and then the number appears on the display board and they [the citizens] have to go to the tables. Then we have this open [area] where our computers and printers are. And people go there, do what they need to do, and if they need help, they will call. This is where I come into the picture" (interview with frontline worker, 2019). In the various municipalities, this new professional role – premised on guiding citizens through digital platforms – has been given different names. In some municipalities, it is labeled a *floorwalker,* in others, a *citizen guide*. However, the premise often remains the same: that people should be taught how to use particular government platforms and that frontline staff should teach them. Such encounters are framed as learning situations, as people should be made capable of doing what they cannot yet do. In some municipalities, frontline workers emphasized that in these encounters, people were required to do it themselves. Indeed, frontline workers would not do it for them, under any circumstances: "We're trying to hold on to this all the time: you must do it yourself. [. . .] We have to forget how many people are in line because . . . Sometimes, it would be easier to say, 'Come on, give me the mouse, then I'll help you.' But we really want them to be able to do it themselves. And then we also need to spend the time it takes" (interview with frontline worker, 2017).

However, while some frontline workers emphasized that they would not take over, there were also a number of situations in which they actually did just that. Not unlike the two modes of ordering found in the tax administration, the situations involved citizens who could not help themselves, and frontline workers would step in. The following fieldnotes illustrate this through a conversation with Maria, who recounts an incident where she had to step in:

> Maria is telling me a story from a case she handled a while back. A woman needed to change her name and get a new passport issued. "I had talked to the Civil Registration System. And at the Civil Registration System office in our town hall, they are usually really helpful," she explains. "But then there are other offices over there [at the town hall] that are more difficult to work with. And in this situation, with this name change, I needed to talk to the Church office. And I almost got into a fight with them. I contacted them several times and . . . And I kept asking them in different ways, because I could sense that, of course, she would be able to change her name. It must be possible. [. . .] So I kept circling around this question, asking the Church office. And he was so stubborn. Then, finally, I got some answers out of him." She pauses a bit in the narrative. "So one of the most challenging things," she then adds, "is when you get into a fight with another part of the governmental branch concerning questions about whether something is possible. Because they're sitting at a desk behind a screen, and we're standing there with a human being. And sometimes we get very passionate about their case, thinking "This cannot be true! There has to be a solution! There has to be something in the system that can help them". [. . .] Maria goes on to say that this particular example is certainly not unique. Indeed, she says that it happens quite frequently that they, the frontline workers, have to negotiate with other branches of the government. "It almost sounds as if," I ask Maria somewhat cautiously, "you become the citizen's lawyer. That you have to carry the person's case with you into the system." Maria nods enthusiastically and agrees, while repeatedly saying yes. "You're actually right. I do not think we have ever thought about it like that up here. But it's often what we do. We dive into the case and try to guide them through a system that, in my opinion, has been digitalized way too quickly." Maria gives a somewhat coy smile. I then ask: "Many of these tasks . . . Shouldn't people do them themselves?" "Well," Maria replies, "Yes, they should." (Fieldnotes, 2017, second author)

In her narrative, Maria goes to great lengths to help the person in question. Not only is she highly engaged in the particular case, but she is also willing to confront other branches of government to provide help. While this kind of explicit conflict rarely surfaced in the conversations with frontline staff, many of its core elements did seem to be shared among them. Frontline workers would often recount how they, out of genuine care and concern for the citizen, would help them more or less directly. In another conversation, talking to a frontline worker about how she would help people in the municipality's computer area, this was made quite explicit:

> **Interviewer:** When you're sitting, then, with them at the computer. Do you then point them to where they should click or do you take the mouse?
>
> **Frontline worker:** Well, that depends. I sort of sense . . . Where are we in terms of. . . [their abilities.] Because, as a starting point, they need to do it themselves. But . . . [Long pause]. But . . . [short pause]. If they say, "Can you please do it for me?" Well, then, I do not stand there and. . . [quarrel]. Then I just take over. And I'm, of course, very thorough when explaining to them how and why I am doing it, and all those things. So . . . But as a starting point, they need to do it themselves.

Recalling a recent series of episodes, the frontline worker talked about how a number of people had needed help with a particular issue: "At the moment, we're at a point where the new trash guidelines have been posted. And then they [the citizens] come and need to get things printed. [. . .] 'Last year, this was printed for me, can this be done again?' [the person will ask]. 'Yes, we can do that' [we reply]. And then we print. Or we say, 'You're welcome to use our computers', and then [they say] 'Well, I cannot.' [Then we say], 'Fine, then I'll do it for you'" (interview with frontline worker, 2019). Something similar applies in the municipalities that provide citizen services through call centers and phones. Frontline workers will attempt to either guide

people over the phone or ask them to come down and receive help at the center. In some cases, however, they might also take over the task and do it for people. "These citizens are often not digital," a frontline worker explained, "'I do not know how to do that' [they will say]. If we, for example, say that they have to go onto our website, that alone will make them say, 'I do not know anything about that!' And in that case you do not, well at least we do not, try anything, then we just say, 'That's fine. Then I'll do it for you'" (interview with frontline worker, 2019).

One of the important things to notice in these encounters where frontline staff take over is that it is often not because they get too tired or frustrated trying to guide people. Instead, they are fueled by the same sense of professional vocation that Maria's narrative exemplifies; they do so because of a genuine wish to help people. While some frontline workers did say that they tried to gently push people to do it themselves, others would underline that if someone said they were unable to perform a task, they would not argue with them.

Complementary and competing modes of ordering

These ethnographic narratives showcase how frontline work takes place within two public sector institutions in Denmark. From these observations, we can begin to see a number of similarities between the institutions in terms of how such work is organized. As already suggested, we propose conceptualizing these patterns as two distinct modes of ordering that are simultaneously contradictory and complementary.

The first mode of ordering, "help to self-help," follows the ideals promoted in the literature on digital era governance as well as in Danish digitalization reforms. This mode of ordering revolves around digital self-service platforms as the main means for citizens to interact with the public sector. The role of frontline workers is mainly to make such interactions possible. Frontline workers must help people become able to help themselves with regard to being able to fill out their taxes, apply for welfare subsidies and otherwise communicate with the public sector through digital means. At both ethnographic sites, this was clearly the primary mode of ordering, i.e. the one that managers spoke about as ideal and the one promoted in official documents and conceptualizations such as no-touch customers. Yet, as we observed, this mode of ordering was not the only one. In certain situations, frontline workers would switch into a second mode; directly helping. Whenever help to self-help failed to actualize its intended outcome, directly helping often took over. This second mode reversed many of the properties of the first. Rather than helping people help themselves, frontline workers solved problems for them. Rather than being promoted, these practices were often left unrecognized or were downplayed in the official policy visions and narratives. And, rather than revolving around an ideal of citizens as self-reliant, the second mode of ordering appeared to be fueled by the frontline workers' professional concern. Table I summarizes these differences.

We suggest that these modes of ordering are both complementary and contradictory. Directly helping is often hidden or underemphasized in official narratives and conceptualizations, only entering when help to self-help is deemed infeasible based on the practical judgment of the frontline worker. Yet, it is precisely the second mode that makes it possible for the first mode to continue and to be maintained on a day-to-day basis. Underneath the ideal of self-sufficiency and helping citizens help themselves, there is an altogether different set of practices that reverse many of the properties of the preferred mode of ordering. Frontline workers act against seemingly universal policy ideals, but in doing so they help maintain and sustain them. Thus, moving continuously between these two modes – switching back and forth – frontline workers are able to truly put digital reforms into action and make them work. This leads to a somewhat paradoxical situation in which the mode of ordering – along with its ideal ways of encountering citizens and organizing bureaucracies – can only be maintained by practices that appear to oppose it. In establishing a form of bureaucracy that has seemingly done away with discretion,

	Mode 1: Help to self-help	Mode 2: Directly helping
Description	Mode of ordering based on welfare staff helping citizens help themselves	Mode of ordering based on welfare staff directly intervening to solve discrete task
Purpose of encounter	To help citizens help themselves (help to self-help)	To help citizens directly, solve particular problems (help)
Problem to be solved	Citizen is unable to solve particular task, citizen must be made capable (citizen as problem)	Citizen is unable to solve particular task, task must be handled by frontline staff (task as problem)
Ideal	Citizen as self-sufficient and self-reliant individual	Welfare work as a care for citizens and their well-being
Practice	Guidance through standardized platforms (hands off)	Help through direct intervention (hands on)
Visibility	Visible in strategies, reports and presented as the official narratives	Visible for frontline workers and the citizens who are being helped

Table I.
Modes of ordering

and leaving all citizens to solve every issue themselves, it is a kind of discretion that is needed to keep things going.

Discussion and concluding remarks

Public administrations and the work that street-level bureaucrats do have become increasingly digitalized over the last two decades (Bovens and Zouridis, 2002; Jansson and Erlingsson, 2014; Buffat, 2015; Pors, 2015; Hansen *et al.*, 2018). Not only has this implied a greater reliance on digital technologies and ICTs in the contact with citizens, it has also warranted a wholesale shift into altogether new organizational forms, as observed by many scholars. Indeed, researchers have suggested that digital era governance is steadily becoming the dominant management paradigm (Dunleavy *et al.*, 2006; Margetts, 2009) and that street-level bureaucracy has turned into system- or screen-level bureaucracy (Bovens and Zouridis, 2002).

With this article, we have sought to shift attention away from high-level descriptions of technological change and bureaucracy, instead offering insights into the day-to-day challenges and dilemmas frontline workers face. Based on cases from two field sites, we have shown how different modes of ordering develop in the interaction between frontline workers and citizens. Formed equally by national reforms, local practices, embodied professional judgments and situational factors, these modes of ordering revolve around making people able to help themselves digitally. Whether by phone or in face-to-face encounters, frontline workers are expected to make people interact with the state in a preferred way: digitally. We have also shown how this preferred mode of ordering relies on a second mode, which is often underemphasized in studies on the digitalization of public administrations and the strategic documents and language used in public sector organizations. In situations where citizens cannot be turned into self-helping individuals, frontline workers switch their practices and help people directly. Rooted in a professional ethos, frontline staff intervene directly in the situation and, at times, become the citizen's voice. This second mode of ordering is vital to sustaining the preferred one, i.e. the first mode, as it makes sure that people otherwise incompatible with the preferred ideals will get the help they need.

These arguments demonstrate that underneath the seemingly linear transition into a new type of digitalized bureaucracy, there are not only tensions that need to be resolved, but a set of underemphasized practices that sustains the everyday order. Rather than decreasing or withering away, then, the discretion of frontline workers takes on altogether new forms. In the seemingly narrow margins for discretion created by official digital policies, new professional

practices emerge, emphasizing the necessity of ethnographic research for understanding the function of public reforms and digitalization. We cannot understand these topics from a distance or by assuming that technological change is inevitable, desirable, or inherently good. Instead, we need to pay close attention to the everyday practices of frontline workers as well as managers – who might be struggling with similar tensions and contradictions between the preferred strategies and what is possible in practice – in public administrations. In this sense, guided by the long-standing efforts of ethnographic researchers, this paper pushes for a greater attentiveness towards the practices entailed in making reforms work – however messy, complicated, and fraught with contradictions they might be.

The paper's arguments also have implications for policymakers and governmental officials. Indeed, highlighting the tensions between different modes of ordering serves to help question the very assumptions of policy ideals. While existing political visions have often relied on a singular image of one ideal mode of ordering – in which bureaucracy is, or is about to become, completely digitalized and the role of frontline workers minimized to helping citizens help themselves – this ideal can only be sustained through the emergence of new work practices. Indeed, our ethnographic work shows how additional modes of ordering are necessary for the preferred modes to function. Our study shows that the seemingly universal policy ideal is challenged and subverted, as it can only be maintained through codified workarounds that cannot be contained in this ideal. This might prompt us to ask policymakers and officials new questions in future research. What would it look like – we might ask – if the ideal of a self-sufficient citizen promoted by policymakers was not claimed to be universal? What if new policy initiatives tried to work generatively with the tensions explored here? Would that better allow the unrecognized work carried out by frontline staff to be seen and acknowledged in new ways? And, if so, what impact might that have on future reforms?

Note

1. All names, places and events mentioned in this article are anonymized.

References

Bhatti, Y., Olsen, A. and Pedersen, L.H. (2011), "Administrative professionals and the diffusion of innovations: the case of citizen service centres", *Public Administration*, Vol. 89 No. 2, pp. 577-594.

Bjerge, B. and Rowe, M. (2017), "Inside the black box of public service change", *Journal of Organizational Ethnography*, Vol. 6 No. 2, pp. 62-67, doi: 10.1108/JOE-05-2017-0023.

Boll, K. (2015), "Deciding on tax evasion – front line discretion and constraints", *Journal of Organizational Ethnography*, Vol. 4 No. 2, pp. 193-207, doi: 10.1108/JOE-07-2014-0018.

Boll, K. and Rhodes, R.A.W. (2015), "Excursions in administrative ethnograph", *Journal of Organizational Ethnography*, Vol. 4 No. 2, doi: 10.1108/JOE-05-2015-0013.

Boswell, J. and Corbett, J. (2015), "Who are we trying to impress?: reflections on navigating political science, ethnography and interpretation", *Journal of Organizational Ethnography*, Vol. 4 No. 2, pp. 223-235, doi: 10.1108/JOE-08-2014-0030.

Bovens, M. and Zouridis, S. (2002), "From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control", *Public Administration Review*, Vol. 62 No. 2, pp. 174-184, doi: 10.1111/0033-3352.00168.

Brodkin, E.Z. (2017), "The ethnographic turn in political science: reflections on the state of the art", *PS - Political Science and Politics*, Vol. 50 No. 1, pp. 131-134, doi: 10.1017/S1049096516002298.

Brodkin, E.Z. and Marston, G. (Eds) (2013), *Work and the Welfare State: Street-Level Organizations and Workfare Politics*, Georgetown University Press, Washington.

Buffat, A. (2015), "Street-level bureaucracy and e-government", *Public Management Review*, Vol. 17 No. 1, pp. 149-161, doi: 10.1080/14719037.2013.771699.

Chadwick, A. and May, C. (2003), "Interaction between states and citizens in the age of the Internet: "e-government" in the United States, Britain, and the European Union", *Governance*, Vol. 16 No. 2, pp. 271-300, doi: 10.1111/1468-0491.00216.

Danish Ministry of Taxation (2006), "Fusionen. Det selvangivelsesløse samfund. afståelsesbegrebet. aktuelle skattepolitiske initiativer", available at: https://www.skm.dk/media/121845/skat-net.pdf.

Dubois, V. (2009), "Towards a critical policy ethnography: lessons from fieldwork on welfare control in France", *Critical Policy Studies*, Vol. 3 No. 2, pp. 221-239, doi: 10.1080/19460170903385684.

Dunleavy, P., Margetts, H., Bastow, S. and Tinkler, J. (2006), *Digital Era Governance: IT Corporations, the State, and E-Government*, Oxford University Press, Oxford.

Elgaard, T.E. (2017), "Materiality, performance and the making of professional identity", *Webb, S, Professional Identity and Social Work*, Routledge, London, pp. 51-62, doi: 10.4324/9781315306957.

European Commission (2019), "The digital economy and society index", available at: https://ec.europa.eu/digital-single-market/desi (accessed 9 October 2019).

Evans, T. (2016), *Professional Discretion in Welfare Services: Beyond Street-Level Bureacracy*, Routledge, London.

Fassin, D. (Ed.) (2015), *At the Heart of the State: The Moral World of Institutions*, Pluto Press, London.

Hansen, H.T., Lundberg, K. and Syltevik, L.J. (2018), "Digitalization, street-level bureaucracy and welfare users' experiences", *Social Policy and Administration*, Vol. 52 No. 1, pp. 67-90, doi: 10.1111/spol.12283.

Henman, P. (2010), *Governing Electronically: E-Government and the Reconfiguration of Public Administration, Policy and Power*, Governing Electronically: E-Government and the Reconfiguration of Public Administration, Policy and Power, Palgrave Macmillan, Basingstoke, doi: 10.1057/9780230248496.

Herzog, L. and Zacka, B. (2017), "Fieldwork in political theory: five arguments for an ethnographic sensibility", *British Journal of Political Science*, pp. 1-22, doi: 10.1017/S0007123416000703.

Hupe, P., Hill, M. and Buffat, A. (Eds), (2016), *Understanding Street-Level Bureaucracy*, Policy Press, Bristol.

Jæger, B. and Löfgren, K. (2010), "The history of the future: changes in Danish E-government strategies 1994–2010", *Information Polity*, Vol. 15 No. 4, pp. 253-269.

Jæger, B. and Pors, A.S. (2017), "Ledelse af digitalisering – fra projekt til præmis", in Aagaard, P. and Agger, A. (Eds), *Ledelse i politisk styrede organisationer*, Hans Reitzels Forlag, Copenhagen, pp. 145-170.

Jansson, G. and Erlingsson, G.Ó. (2014), "More E-government, less street-level bureaucracy? On legitimacy and the human side of public administration", *Journal of Information Technology and Politics*, Vol. 11 No. 3, pp. 291-308, doi: 10.1080/19331681.2014.908155.

Law, J. (1994), *Organizing Modernity*, Blackwell, Oxford.

Lipsky, M. (2010), *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service (30th Ann.)*, Russell Sage Foundation, New York, NY.

Margetts, H. (2009), "Public management change and E-government: the emergence of digital era governance", in Chadwick, A. and Howard, P. (Eds), *Routledge Handbook of Internet Politics*, Routledge, London, pp. 114-128.

Maynard-Moody, S. and Musheno, M. (2000), "State agent or citizen agent: two narratives of discretion", *Journal of Public Administration Research and Theory*, Vol. 10 No. 2, pp. 329-358.

OECD (2014), *Increasing Taxpayers' Use of Self-Service Channels*, OECD Publishing, Paris, doi: 10.1787/9789264223288-en.

Pollitt, C. (2011), "Mainstreaming technological change in the study of public management", *Public Policy and Administration*, Vol. 26 No. 4, pp. 377-397, doi: 10.1177/0952076710378548.

Pors, A.S. (2015), "Becoming digital – passages to service in the digitized bureaucracy", *Journal of Organizational Ethnography*, Vol. 4 No. 2, pp. 177-192, doi: 10.1108/JOE-08-2014-0031.

Rhodes, R.A.W. (2014), "Genre Blurring and public administration: what can we learn from ethnography?", *Australian Journal of Public Administration*, Vol. 73 No. 3, pp. 317-330, doi: 10.1111/1467-8500.12085.

Schou, J. and Hjelholt, M. (2018), *Digitalization and Public Sector Transformations*, Palgrave Macmillan.

The Government, Local Government Denmark and Danish Regions (2016), *A Stronger and More Secure Digital Denmark: Digital Strategy 2016–2020*, Danish Agency for Digitization, Copenhagen.

Tummers, L.L., Bekkers, V., Vink, E. and Musheno, M. (2015), "Coping during public service delivery: a conceptualization and systematic review of the literature", *Journal of Public Administration Research and Theory*, Vol. 25 No. 4, pp. 1099-1126, doi: 10.1093/jopart/muu056.

West, D.M. (2005), *Digital Government: Technology and Public Sector Performance*, Princeton University Press, Princeton.

Zacka, B. (2017), *When the State Meets the Street: Public Service and Moral Agency*, Harvard University Press, London.

Corresponding author
Bastian Jørgensen can be contacted at: bmjo@itu.dk

Research paper 2:

Organizing artificial intelligence - Representing work in the Danish Customs and Tax Administration

Jørgensen, B., Gad, C., & Winthereik, B. R. (in press). Organizing Artificial Intelligence. In D. Lanzeni, S. Pink & K. Waltorp (Eds.), *Contested Technologies: Future Anthropologies*. Bloomsbury.

Organizing artificial intelligence

Representing work in the Danish Customs and Tax Administration

Keywords: Artificial Intelligence (AI), machine learning algorithms, public sector digitalization, data work, ethnography.

Bastian Jørgensen, Christopher Gad, Brit Ross Winthereik

IT University of Copenhagen

Introduction

The allure of advanced techniques for data analytics pervades contemporary imaginaries. As such Artificial Intelligence (AI) has been recognized as a globally important topic in reports published by the European Commission (EC 2018; EC 2020). According to these reports AI already influences the daily affairs of European member states and their citizens, and they pose that the total global spending on AI will only increase in the years to come.

The Danish government has advocated that the country should aim to become a frontrunner in AI through an ethical and responsible perspective on algorithms (Danish Government 2019). This ambition shows how AI has become a site for contestation over how it is developed, implemented and used. Controversies around algorithms that have led to responses like the report by the Danish government, are rooted in calls for recognizing embedded authoritarian tendencies, and it has been argued that AI tends to obfuscate decision-making processes with dire consequences for people's lives (Danaher 2016; Pasquale 2015; Zuboff 2019). Both proponents and critics seem to agree that advanced techniques for data analytics are already pervasive in society and carry the potential to radically change organizations and societies. For both camps the question is not, whether or how AI works, but how to recognize the potentials/dangers of emerging technologies.

This is where anthropology can step in, because what if the problem is not about what kind of AI to develop (authoritarian or ethical)? What if the problem is something entirely different and concerns whether or how we invest in the people and the work practices that are to 'receive' AI?

In response to the growing advocacy for using advanced analytics, researchers have called for ethnographic studies of the use of algorithms and data (Iliadis and Russo 2016; Kitchin 2017; Seaver 2017). Hitherto, much critical scholarship has namely focused on strategies and behind-the-scenes analyses of big tech companies such as Amazon, Google and Facebook (Pasquale 2015; Zuboff 2019), or on new forms of making business in organizations established 'after the internet' (Seaver 2018; Geiger 2017). Less attention has been paid to mundane practices around the development and use of artificial intelligence, and even fewer studies focus on AI in the public sector (for a notable exception see Reutter and Spilker 2019).

This paper analyzes a recent attempt to develop and implement advanced analytics in the Danish Customs and Tax Administration. The paper builds on scholarship within anthropological studies of technology (Suchman 1987, 1995, 2000, 2007; Seaver 2017, 2018; Forsythe 1993a, 1993b), Science and Technology Studies (STS) (Berg 1997; Star 1995; Star and Strauss 1999; Zuboff 2019), and work broadly conceived as critical data studies (Rieder and Simon 2016; Selbst and Barocas 2018; Iliadis and Russo 2016). It shows how the development of a new machine learning algorithm happened in conversation with already existing infrastructures and divisions of work (Karasti et. al. 2016: 7), and argues that while algorithms depend on existing forms of organization, we should also make sure to attend to the arrival of new categories. One example of such a new category, which our analysis emphasizes, is *data work*.

The empirical material on which the analysis rests is generated in The Risk Score Project in The Danish Customs and Tax Administration. The aim of this project was to develop a machine learning algorithm to assist the conduct of customs inspection. The hope was that the unit in charge of the project could use AI to develop a more efficient method for identifying packages imported to Denmark that might not have been taxed properly, and then mark those packages for further inspection. In the chapter we will draw attention to how the development of a machine learning algorithms is based on separating what we refer to as *data work* from the other forms of work that the algorithms are imagined to support. The aim of the chapter is twofold. On the one hand, it is to show what can happen when organizational contexts and their forms of contestation are disregarded in the development of digital technologies. On the other hand, we wish to question the widespread idea that advanced analytics constitute an unquestionable improvement of existing forms of work, even the forms of work classified as 'routine work' (Suchman 2000).

Towards a Data-driven Tax Administration

In recent years, the Danish Customs and Tax Administration has invested heavily in building its expertise and capacity for developing advanced analytics. In 2014 the tax administration established a new office which was referred to as the 'Centre of Excellence for Advanced Analytics and Machine Learning'. By 2018 the office had engaged 35 full-time employees including 27 data scientists. This development overlap with, the tax administration as a whole undergoing a significant reorganization process. In 2018, The Danish Government split the tax administration into seven independent agencies and made those agencies responsible for managing different areas of expertise. One of the new agencies was named the IT and Development Agency. With HQs in the Copenhagen Capital Region and about 1000 employees across approximately 70 offices nationwide,[1] this agency became the 3rd largest of the 7 new agencies. This development reflects the current political and managerial belief in the importance of investing in and developing the *internal* IT capacities of the State, as expressed in a government report which introduced the newly organized tax administration: "New organizing, more employees, and IT that works" (The Ministry of Taxation 2016, our emphasis).

The organizational changes were in line with technologically optimistic ideas expressed by the OECD's Forum on Tax Administration (FTA) (OECD 2019). In a series of reports called 'Tax Administrations of the Future Series' the OECD urges tax administrations to stimulate the development of 'a data-driven culture' (OECD 2016a; OECD 2016b; OECD 2016c). The report 'Technologies for Better Tax Administration' states:

> Bringing a data-driven culture into a tax administration implies developing a culture in which data and analysis drive all aspects of the organization. This is a journey, not a destination. It will continue as new technologies emerge; digital disruption creates even greater innovation and customer expectations change (OECD 2016c: 58).

This statement combines technological determinism (data as the driver) with a processual outlook (an imagery of the future where the only stable element seems to be continuous emergence of new technologies). The vision is clear: Tax administrations must continually transform themselves in order to stay relevant in the 21st century. In the report 'Advanced Analytics for Better Tax Administration – Putting Data to Work' attention is directed towards how tax administration can "extract value from data using advanced analytics" (OECD 2016a: 3). The report uses a cross-national survey to show that the primary use of advanced analytics is related to inform case selection, achieve tax compliance and manage taxpayers' debts (OECD 2016a: 27). As the report states it "makes no assessment of the relevant capability of administrations working in these areas; it seeks only to identify where work is being carried out". Rather than discussing specific technologies, implementation strategies or challenges. The report animates a mapping of where work that could be supplemented or replaced by AI is happening. As we see in figure 1, European tax administrations have many areas where AI could potentially play a role.

Table 2.1. **Summary of activities by country**

	Audit case selection	Filing & payment compliance	Taxpayer Service	Debt Management	Policy
Australia	■	■	■	■	■
Canada	■	■	■	■	
China		■			■
Finland	■	■		■	■
France	■				
Ireland	■	■	■	■	■
Malaysia	■			■	
Mexico	■	■			■
Netherlands	■			■	
New Zealand	■	■	■	■	■
Norway	■	■	■	■	
Singapore	■	■	■	■	■
Sweden	■			■	■
Switzerland	■				
United Kingdom	■	■	■	■	■
United States	■	■	■	■	■

Source: FTA Advanced Analytics survey, 2015.

Figure 1 – Activities of Tax Administrations within Advanced Analytics (OECD 2016a: 20)

In what follows, we analyze some of the concrete outcomes of these imaginaries in The Danish Customs and Tax Administration.

Fieldwork

In February 2017, shortly after first author commenced his fieldwork in the office for advanced analytics and machine learning, a manager introduced The Risk Score Project. The goal of the project was to develop a new method for determining risk scores, which assist the customs department in detecting tax and VAT fraud for packages imported to Denmark. First author and the manager agreed that we would be allowed to follow the development of the new risk score model based on machine learning principles. Soon after, however, the future of the project became uncertain, when a newly appointed manager explained that he had been employed to close down the project. First author was advised to study 'something else'. He followed this advice, but remained in contact with the participants in and around developing algorithms for risk scoring. This was possible, because even though The Risk Score Project was terminated, the data scientists kept working to deliver an algorithm, which could be tested on the custom officers' inspection of packages.

Empirically, the present chapter draws on observations, interviews, document studies, and email conversations with the data scientists and their collaborators in developing and testing the algorithm for inspection (see Fig. 2).

Informants:	**Office:**	**Location(s):**
Data Scientists	Advanced Analytics	Northern Copenhagen
System Monitors	The Monitoring Unit	Southern Copenhagen
Customs Officers	Customs offices	E.g. borders and airports

Figure 2 – Overview of informants' work titles and work places

The empirical material was generated from February 2017 to July 2018.

Theoretical-analytical resources for situating AI

The belief that artificial intelligence (AI) has the potential to improve private and public organizations is not unprecedented. As M.C. Elish and danah boyd notes, a similar enthusiasm characterizing discussions around AI in the 1980s and 1990s. In this period, a strong interest in AI, in the form of expert systems, spread from universities to commercial settings (Elish and boyd 2018: 4). If we look back at some of the research that happened in the 1980s, we find computer scientists who were developing systems that could imitate human experts. Expert systems, as they were called, were widely imagined to be capable of replacing crucial elements of an expert's work, and the computer scientists were tasked with explicating the rules and facts that experts would supposedly follow in their daily work. 'Facts' were stored in databases often referred to as 'the knowledge base', whereas rules for how experts would apply facts, were used as the foundations for an 'inference mechanism' (see Fig. 3). By embedding facts and rules in code. expert systems were imagined to, for example, be able to provide relevant information to help doctors diagnosing patients.

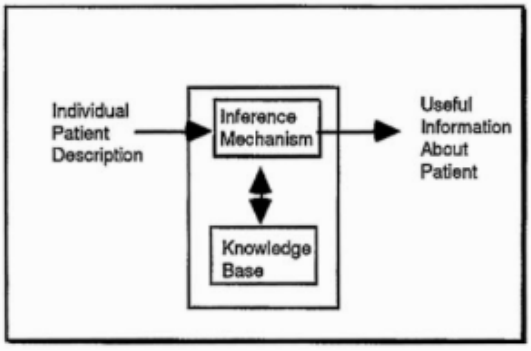


Figure 3 - Conceptual Design of an Expert System (Berg 1997)

Assisting doctors in diagnosing patients was one of the most popular cases for research and development of expert systems. An important assumption at stake was that automating certain parts of the doctors' work, would relieve them of time-consuming routine tasks and 'rationalize' their work by eliminating personal idiosyncrasies in doctors' work practices (Berg 1997).

Since experts carry out many tasks that cannot be reduced to rules or formulae without a significant loss of meaning, it was soon discovered that expert systems would never be able to replace absolutely crucial parts of expert work. And so, expert systems never really made it into the thick of clinical work practices. One of the reasons that was pointed out was the problem of idealization of work and the necessity of disregarding social and organizational contexts (Collins 1987; see also Collins 2018).

As an anthropologist working in the tech industry of 'the Valley', Lucy Suchman proposed instead to focus on how facts and rules would emerge as part of situated action (Suchman 1987). Through a landmark study of some of the world's brightest computer scientists failing to operate a 'smart' photo copier, her PhD dissertation argues that not only is intelligent action an emergent phenomenon, expertise must also be understood in relation to the material and social circumstances, in which experts act (ibid.: 70). Intelligence depends on contextual resources available, but as such resources cannot easily be formalized, they remain invisible (see also Star and Strauss 1999). Suchman ultimately warned against operating with a narrow view of intelligence as symbol manipulation, and not to assume similarity between humans' activity and machine operations when developing automated systems (Suchman 2007: 37; see also Broussard 2018).

Fast forward to today, where artificial intelligence seems to have shifted from being concerned with expert systems to being concerned with machine learning. Research describes this as a shift from a *rule based logic* to a *machine learning based logic* (de Vries 2013), from a *top-down approach* to a *bottom-up approach* (Dourish 2016: 7), and from *symbolic representation of human knowledge and procedural logic* to the *crunching of vast amounts of data, detecting patterns and producing probabilistic results* (Elish and boyd 2018). The shift is very important for understanding the role of work and expertise in AI projects today.

More specifically, constructing a 'knowledge base' for explication of 'facts' has been replaced by methods to create facts by means of data in a database. The assumption now is that knowledge is already present in the data; it just needs a bit of cleaning, sorting, and analysis. Rather than computer scientists explicitly constructing and defining the rules for the 'inference engine', machine learning is based on the idea that the algorithm will itself infer rules by identifying statistical patterns in historical data. Data scientists will then develop models for presentation of those patterns (some of those are referred to as 'dashboards').

This embedding of expert knowledge may be one of the reasons why current public debates of AI tend to focus on the opaqueness of machine learning algorithms (see also Burrell 2016; Castelvecchi 2016).[2] Still, despite attempts to develop 'ethical AI', what happens 'on the ground' is that machine learning algorithms are developed that embed the idea that there is no need to explicate the work and knowledge of experts, because the knowledge is in the data. This has organizational and societal implications. In the following section, we dive into specific ramifications of such developments in The Danish Customs and Tax Administration. We follow a suggestion put forward by Elish and boyd to situate the magic of AI (2018).[3] We also follow Ziewitz, who asks: What do we mean when we talk about algorithms? (2016). Our aim is not to 'disenchant' AI, but to scrutinize relations between AI and existing practices and infrastructures in a public sector organization.

Work in the Customs Inspection

The process of selecting packages for customs inspection is already highly reliant on IT systems and algorithms. The current risk system is managed by system monitors, and assesses packages using three different methods: *Risk profiles*, *risk scores*, and *random selection*. As indicated by the name of The Risk Score Project the algorithm to be developed was meant to replace just one of the existing ways of calculating risk (*risk scores)*, which marks packages for inspection based on variables such as weight, price, country of origin and the package type (The National Audit Office of Denmark 2017: 22). In the future, if a risk score for a package exceeds a certain threshold, defined by system monitors and customs officers, the risk system will flag the package for human inspection. The idea of replacing the existing rule-based risk score with a machine learning based risk score, can be compared to the shift from expert systems to machine learning as outlined above.

At the time when the project commenced, replacing the existing risk score calculation with a machine learning based risk score seemed like an obvious choice. This was first of all the case, because the current risk score had not been in use since 2014, as reported by The National Audit Office of Denmark:

> The Danish Customs and Tax Administration has given up on this part of the risk system and is now developing a new function instead. This new risk score will be created automatically based on

> historical data and the results from previous inspections. It will provide the tax administration with the opportunity to find abnormal patterns, which can cause an inspection. The tax administration does not know yet when the new risk score can be taken into use (The National Audit Office of Denmark, 2017: 23).

Replacement was also in line with the OECD report mentioned above which recommends using advanced analytics for 'case selection' in order to *put data to work* (OECD 2016a).

During a workshop in the system monitors unit first author observes some of the discussions around implementing machine learning algorithms. About ten employees from the Danish Customs and Tax Administration are present.

> "A large table divides the employees into two groups: Four system monitors are seated at one of the sides and three data scientists on the other. At the end of the table, a project manager stands next to a whiteboard and guides a brainstorm. Within an hour, the project manager has attached thirty post-it notes to the whiteboard. On the post-its, I read sentences like: *"More data", "A better understanding of data", "Different types of algorithms"*"" (Except from fieldnotes, Jørgensen).

The conversations at the workshop concerned questions around what data would be available for the algorithm and what algorithm to deploy. A data scientist told that he had spent 90% of his time in the risks score project "preparing data" and "establishing an environment" for analysis, and only 5% of his time on "actually writing code." The workshop illustrated some of the work that goes into implementing an algorithm: Bringing different professional groups into the same room, facilitation of a discussion, as well as sense making around whether the current conditions in the organization allows for the technology. First author was puzzled that no customs officers were present as participants. When he asked a data scientist about the absence of the customs officers at the workshop, the data scientist told him that the system monitors were 'the customers' of the project. The logic was that system monitors were the ones who managed the current risk score, and they were the ones who were knowledgeable about IT systems and data, not the customs officers. The attention of the data scientist

was thus directed towards understanding IT systems and databases rather than understanding the work practices of the customs officers.

Is this a case of bad user involvement? Maybe, but something else may also be at play: What is at stake is a representation of work in the tax administration that favors *data work* over practical, situated work that the algorithm is supposed to support. In the remaining part of this section, we further explore the enacted distance between data work and other kinds of work.

While customs officers were not present in the workshop, one question at the workshop directed everyone's attention towards their work. Just before the end of the brainstorm, one of the system monitors raised his hand and said: "I just thought about an additional need. It would be beneficial if the customs officers could get an explanation of why the algorithm selects a specific package for inspection." After a moment of silence, one of the data scientists replied: "The problem with some of these algorithms is that they are a bit black-boxed. That means that they are difficult, if not impossible, to explain. But we are testing another algorithm called LIME, which can provide explanations of machine learning algorithms."

The system monitors raised a concern about how the data work happening at the workshop would align with the practical work of customs officers. But rather than making further inquiry into what kinds of explanation the customs officers might need and why – questions that would direct attention further towards the practical work of the customs officers – the response by the data scientists immediately redirected attention back to data and to the development and implementation of explanation algorithms.

Later, when asked about the importance of getting an explanation a system monitor tapped his nose with his finger and said: "The question is how we make sure that the nose of the customs officers become part of the system." He elaborated that for customs officers to perform an efficient inspection, they would have to know both why and how to inspect a given package. The new risk score seemed prone to embed its own risks simply because it would be difficult to and time-consuming for customs officers to figure out how to inspect packages without an explanation of why the package had been scored as suspicious in the first place.

In the article 'How to Talk About the Body', Bruno Latour (2004) argues that bodies learn to be affected by material surroundings and conventions (Latour 2004: 205). He provides an example of how specialists ('noses') are trained in the perfume industry. The training of 'noses' happens through the use of an

odour kit, which teaches the 'noses' to discriminate subtle differences in smells and to tell different smells apart from one another. If we compare to what was said in the workshop, for the system monitor, the nose of the customs officers could be an important 'addition' to the algorithm. Yet, the trouble was that the customs officers' 'noses' had already been through a de-skilling process due to years of systematic attention and funding of digitalization and data work.

> **Customs officer:** [In the past] we would get a large flight manifest with information about what kind of goods there was on board of airplanes. Who is the sender, who is the recipient, and what is the package declaration? Then we could see 'ahh a plastic bunny for Mr. Jensen; there is always something shady about that.' How do we get the system to think like that? [Today] we can ask for them [the flight manifests], but they always arrive at night, and we do not have night shifts anymore, so most of the time, we do not look at them. If we went through the manifests again, it would require that we were allowed to have night shifts.

The customs officer's reflections on how to use his 'nose' in this interview does not focus on knowing why packages are selected for inspections in advance. Instead, the customs officer focused on the use of materials, such as flight manifests, and former ways of organizing work, such as night shifts. In this way, the customs officer situated *the nose* as part of the material learning practices and organization of work. The customs officer insinuated that *the nose* had already been lost due to earlier organizational changes which introduced the division of labor into the selection of packages and inspection. This had removed customs officers from doing an essential part of customs inspection –which allowed their *nose* to be trained (cf. Latour 2004). In this light the new machine learning algorithm would only seem to reproduce and fortify this situation. According to another customs officer, the turning point was the formation of the monitoring unit in 2014. When asked if system monitors were able to represent their work in The Risk Score Project a customs officer stated:

> **Customs officer:** The system monitors do not have their 'hands in the dirt'. They don't look at the result of inspections. That was exactly what was so great back in the days where we decided which

> packages to inspect by looking at the flight manifests. It was my job to determine which packages to inspect and to perform the actual inspection; in that way, I became smarter and smarter. I learned that a company was like this and that and that we should not inspect packages from that company. We became smarter and smarter from our inspections. We still inspect, but where does our knowledge go?

From this perspective, the nose of the customs officer is not just an expression of a specific type of human intelligence or tacit knowledge, which customs officers possess, and which can be trained over time. Rather it is expertise that is contextual and situated, and maintained through interaction and relations to materials (flight schedules), ways of organizing (routines, relations to other processes). A strict separation of work processes into the tasks of selecting and inspecting packages seems to put the contextual knowledge in danger and possibly create new kinds of risk, because it is only the IT-system, not the human worker that will learn from experience.

Discussion

As public sector organizations direct attention towards artificial intelligence (AI), and in particular the application of machine learning algorithms, they risk losing sight of the practical work they are imagined to support. Above we have shown how machine learning algorithms work on the basis of historical data, rather than on input from human experts. The algorithm moreover requires a work situation where data, in one way or the other, is already present or that data from outside of the organization can be procured. The risk score project exemplifies how processes that human experts consider part of the work practice are divided up. Thus, the development of a machine learning algorithm creates distances in the formal organization of work. The distance is exacerbated when system monitors, who manages the existing IT-systems, and customs officers, who inspects packages, are not brought into dialogue as part of the development process.

Because the development of machine learning algorithms in the public sector involves organizational units that are not already collaborating, it may not be straightforward for management, who should actually be part of the development of a technology that is imagined to be working autonomously, but which potentially impacts the organisation across several units. Viewed as a tool for organizations to make more accurate and

efficient decisions, machine learning algorithms direct anthropological attention towards the relation between emerging technologies and the wider situation of work.

In The Risk Score Project the machine learning algorithm was imagined to assist the overall customs inspection by being one among several methods for selecting packages to be inspected. Following Suchman (2000, 2007) we can describe this as directing attention to the knowledge work of an organization at risk of becoming cast in terms of routine work. Representing work cast as routine work in terms knowledge work requires ethnographic insight into the day-to-day organization and discussions in a project, as well as to how 'the project' is itself an organizing frame that directs resources to some activities and not others.

When organizations invest in improving knowledge work, they may then assume that current organizational inefficiencies are caused by a lack of knowledge. Studying work in the Danish Customs and Tax Administration, made it clear that sometimes inefficiencies and risks are distributed and enacted in new ways through the automation of a small part of a complex process. In this way, it seemed as if a more accurate logic for selecting package would, at best, only be able to make minor improvements to the overall customs inspection. One may ponder – given the difficulties demonstrated here – how public sector institutions assess the potential risks when implementing AI in work practices that are themselves emergent like in the social services, for example.

Conclusion

We have argued that the development of machine learning algorithms is connected to changes in what counts as human intelligence and expertise. We have brought forth the concept of the custom officer's nose as a way of highlighting how human intelligence in a machine learning project was framed as 'additional knowledge' that could be embedded in the algorithm. One reason for why anthropologists should be concerned with emerging technologies is that an emerging technology like a machine learning algorithm forms a window into human-machine relations where human labor is represented and configured in particular ways, which makes certain kinds of expertise obsolete. The introduction of machine learning in the Danish Customs and Tax Administration illustrates that expertise is varied, but also that some forms are harder to represent. Since the customs officers were not the customers of The Risk Score Project, but 'just' users, we may ask if the

orientation towards user-involvement in technology development that anthropologists have successfully inserted themselves into, has transitioned into a new phase. If so, anthropologists must equip themselves to be partners in and analysts of user-driven tech-development without users.

Notes

[1] See https://www.ufst.dk/media/2645/ufstdk-ufst-org-diagram_090819.pdf.

[2] Accentuating the fact that understanding IT-systems, code and algorithms in general require expert knowledge, machine learning principles have been conceptualized as complete black boxes. Legal scholars have discussed the implication of use of black-boxed algorithms for citizens and for the conduct of law, and suggested that special legislation is required (Goodman and Flaxman 2017; Selbst and Barocas 2018). This has also led to the development of a subfield in computer science on so-called 'explanation algorithms' - algorithms able to provide 'explanations' of those algorithms considered black boxes (Ribeiro et. al. 2016; Adadi and Berrada 2018). An anthropological response to this is forming at University of Copenhagen (SODAS), where Morten Axel Pedersen calls for a 'machine anthropology'.

[3] "In the face of an increasingly widespread blind faith in data-driven technologies, we argue for grounding machine learning-based practices and untethering them from hype and fear cycles" (Elish and boyd 2018).

References

Adadi, A., and M. Berrada (2018). 'Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)', *IEEE Access,* 6: 52138-52160.

Berg, M. (1997). *Rationalizing medical work: decision-support techniques and medical practices*. Cambridge, MA: MIT press.

Broussard, M. (2018) *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press.

Burrell, J. (2016) 'How the machine 'thinks': Understanding opacity in machine learning algorithms', *Big Data & Society*. https://doi.org/10.1177/2053951715622512.

Castelvecchi, D. (2016) 'Can we open the black box of AI?', *Nature News,* 538(7623): 20.

Collins, H. M. (1987) “Expert systems and the science of knowledge.” In *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology,* edited by Wiebe E. Bijker, Thomas P. Hughes, and Trevor J. Pinch, 329-348. London: MIT Press.

Collins, Harry M. (2018) *Artifictional Intelligence: Against Humanity's Surrender to Computers*. Cambridge, MA: MIT Press.

Danaher, J. (2016). ‘The threat of algocracy: Reality, resistance and accommodation’, *Philosophy & Technology,* 29(3): 245-268.

Danish Government. (2019). *National strategi for kunstig intelligens*. Accessed February 3, 2020. https://www.regeringen.dk/media/6537/ai-strategi_web.pdf.

de Vries, Katja. 2013. ‘Privacy, due process and the computational turn: A parable and a first analysis.’ In *Privacy, Due Process and the Computational Turn*, edited by Mireille Hildebrandt and Katja de Vries. New York: Routledge.

Dourish, P. (2016). ‘Algorithms and their others: Algorithmic culture in context’, *Big Data & Society*. https://doi.org/10.1177/2053951716665128

Elish, M. C., and d. boyd (2018). ‘Situating methods in the magic of Big Data and AI’, *Communication Monographs, 85*(1): 57-80.

European Commission. (2018). *Artificial Intelligence for Europe*. Brussels. Accessed February 3, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN

European Commission. (2020). On *Artificial Intelligence – A European approach to excellence and trust*. Brussels. Accessed February 3, 2020. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

Forsythe, D. E. (1993a). ‘Engineering knowledge: The construction of knowledge in artificial intelligence.’ *Social Studies of Science,* 23(3): 445-477.

Forsythe, D. E. (1993b). ‘The construction of work in artificial intelligence’, *Science, Technology, & Human Values,* 18(4): 460-479.

Geiger, R. S. (2017). 'Beyond opening up the black box: Investigating the role of algorithmic systems in Wikipedian organizational culture', *Big Data & Society,* 4(2). https://doi.org/10.1177/2053951717730735.

Goodman, B., and S. Flaxman. (2017.) 'European Union regulations on algorithmic decision-making and a "right to explanation"', *AI Magazine,* 38(3): 50-57.

Iliadis, A., and F. Russo. (2016). 'Critical data studies: An introduction', *Big Data & Society, 3*(2). https://doi.org/10.1177/2053951716674238.

Karasti, H., F. Millerand, C. M. Hine, and G. C. Bowker. (2016). 'Special Issue: Knowledge infrastructures: part I', *Science & Technology Studies*.

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

Latour, B. (2004). 'How to talk about the body? The normative dimension of science studies', *Body & society, 10*(2-3): 205-229.

Ministry of Taxation. (2016). *Et nyt skattevæsen – Ny organisering, flere medarbejdere og velfungerende IT*. Accessed February 3, 2020. https://www.skm.dk/media/1371591/Et-nyt-skattevaesen.pdf.

National Audit Office of Denmark. (2017). *Rigsrevisionens beretning om SKATs kontrol og vejledning på toldområdet.* Accessed February 3, 2020. https://www.rigsrevisionen.dk/media/2104734/sr0717.pdf.

OECD. (2016a). *Advanced Analytics for Better Tax Administration: Putting Data to Work*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264256453-en.

OECD. (2016b). *Rethinking Tax Services: The Changing Role of Tax Service Providers in SME Tax Compliance.* Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264256200-en.

OECD. (2016c). *Technologies for Better Tax Administration: A Practical Guide for Revenue Bodies*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264256439-en.

OECD. (2019). 'Forum on Tax Administration.' Accessed February 3, 2020. https://www.oecd.org/tax/forum-on-tax-administration/.

Pasquale, F. (2015). *The Black Box Society*. Cambridge, MA: Harvard University Press.

Reutter, L. M., and H. S. Spilker. (2019). 'The Quest for Workable Data - Building Machine Learning Algorithms from Public Sector Archives.' In *The Democratization of Artificial Intelligence - Net Politics in the Era of Learning Algorithms*, edited by Andreas Sudmann. Bielefeld: Transcript Verlag.

Ribeiro, M. T., S. Singh, and C. Guestrin. (2016). 'Why should I trust you?: Explaining the predictions of any classifier.' In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, ACM*: 1135-1144.

Rieder, G., and J. Simon. (2016). 'Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data', *Big Data & Society*, 3(1), https://doi.org/10.1177/2053951716649398.

Seaver, N. (2017). 'Algorithms as culture: Some tactics for the ethnography of algorithmic systems', *Big Data & Society*, 4(2). https://doi.org/10.1177/2053951717738104.

Seaver, N. (2018). 'What should an anthropology of algorithms do?', *Cultural Anthropology,* 33(3): 375-385.

Selbst, A. D., and S. Barocas. (2018.) 'The intuitive appeal of explainable machines', *87 Fordham Law Review 1085*.

Star, S. L., ed (1995). 'The politics of formal representations: Wizards, gurus, and organizational complexity.' In *Ecologies of Knowledge: Work and Politics in Science and Technology*, 88-118. Albany: State University of New York Press.

Star, S. L., and A. Strauss. (1999). 'Layers of silence, arenas of voice: The ecology of visible and invisible work', *Computer Supported Cooperative Work (CSCW)* 8(1-2): 9-30.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Suchman, L. (1995). 'Making work visible', *Communications of the ACM 38*(9): 56-64.

Suchman, L. (2000). 'Making a case: 'knowledge' and 'routine' work in document production.' In *Workplace studies: Recovering Work Practice and Informing System Design*, edited by Christian Heath, Jon Hindmarsch, and Paul Luff, 29-45. Cambridge: Cambridge University Press.

Suchman, L. (2007). *Human-machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press.

Ziewitz, M. (2016). 'Governing Algorithms: Myth, Mess, and Methods', *Science, Technology, & Human Values, 41*(1): 3–16. https://doi.org/10.1177/0162243915608948

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. London: Profile Books.

Research paper 3:
Bringing data home: The reconfiguration of public data infrastructures and public-private boundaries

Jørgensen, B. (2021). Bringing data home: The reconfiguration of public data infrastructures and public-private boundaries [Manuscript submitted for publication].

Bringing data home: The reconfiguration of public data infrastructures and public-private boundaries

Journal:	*Big Data & Society*
Manuscript Type:	Original Research Article
Keywords:	data infrastructures, reconfiguration, public-private relationship, data frictions, data politics, digital welfare state
Abstract:	It is often argued that a consequence of the data revolution is that the boundaries between public and private actors become increasingly blurred and intermeshed. Several scholars have discussed how the public sector's investment in data infrastructures can lead to a growing dependency on private companies and the privatization of public assets and services. This article explores the changing relationship between public and private actors based on an ethnographic study of the development of a data infrastructure in the Danish Customs and Tax Administration. It explores organizational frictions in the tax administration's attempt to take back responsibility and control over its data infrastructures. Drawing on Lucy Suchman's notion of reconfiguration, the article explores the figure of 'bringing home', which has become prevalent in the Danish public sector. The figure of 'bringing home' is a response and an alternative vision to several decades of outsourcing the development and management of Government IT systems, including the responsibility for sharing tax data. The paper shows how the figure of 'bringing home' becomes translated and transformed as public employees work to develop a new data-sharing infrastructure. The paper contributes with insights into the construction of public data infrastructures. It discusses how the boundaries between public and private actors become reconfigured through in the development of public data infrastructures. Reconfiguring the boundaries between public and private actors is a continuous effort in the public sector, and it is an important task for organizational and political studies of data infrastructures to discuss these changes.

Bringing data home: The reconfiguration of public data infrastructures and public-private boundaries

Abstract

It is often argued that a consequence of the data revolution is that the boundaries between public and private actors become increasingly blurred and intermeshed. Several scholars have discussed how the public sector's investment in data infrastructures can lead to a growing dependency on private companies and to the privatization of public assets and services. This article explores the changing relationship between public and private actors based on an ethnographic study of the development of a data infrastructure in the Danish Customs and Tax Administration. It explores organizational frictions in the tax administration's attempt to take back responsibility and control of its data infrastructures. Drawing on Lucy Suchman's notion of reconfiguration, the article explores the figure of 'bringing home', which has become prevalent in the Danish public sector. The figure of 'bringing home' is a response and an alternative vision to several decades of outsourcing the development and management of Government IT systems, including the responsibility for sharing tax data. The paper shows how the figure of 'bringing home' becomes translated and transformed as public employees work to develop a new data sharing infrastructure. The paper contributes with insights into the construction of public data infrastructures and discuss how the boundaries between public and private actors become reconfigured through the work and decisions of public employees. Reconfiguring the boundaries between public and private actors is a continuous effort in the public sector that should be open to scrutiny, contestation, and discussion.

Keywords: Data infrastructures, reconfiguration, public-private relationship, data frictions, data politics, the digital welfare state

Introduction - The data revolution and its frictions

The last decade has been characterized by ongoing hype concerning the potentials and promises of data (Kitchin, 2014; Hockenhull & Cohn, 2020). Data, whether 'big', 'small', 'thick', or 'thin', have become a gravitational point around which organizations re-imagine themselves (Heaphy, 2019; Hong et al., 2019; Stalph, 2020). Indeed, data have been imagined as providing organizations with unseen capacities, such as analytical precision in prediction and new scope for managers' decision-making power. While experimentation with data began in the private sector public sector organizations have followed suit. The discourse of better control and more efficient usage of data promises to make public organizations more efficient and enable them to deliver better services to citizens in times of austerity. Government authorities are, for example, testing advanced algorithms to help detect terror tactics, crime, and fraud (Amoore, 2009; Christin, 2017; Benbouzid, 2019), and with making government data open and accessible to private companies and the general public. The rationale is that open data can contribute to a more transparent public sector and provide the basis for private businesses to create new products and services, thereby fostering economic growth (Bates, 2012; Kitchin, 2014; Ratner & Ruppert, 2019). The extent to which governments have realized these promises (Høyer, 2019) and expectations (Fiore-Gartland & Neff, 2015) of data is an open question. Still, it is beyond any doubt that 'data talk' has taken a firm hold of public organizations as a language with a strong performative effect (Bigo et al., 2019).

As a response to the many and varied effects of the proliferation of data talk, critical scholars have called for empirical investigations of data practices and politics (Kitchin & Lauriault, 2014; Bigo et al., 2019). Sociologists Didier Bigo, Engin Isin, and Evelyn Ruppert, for instance, have proposed that scholars direct their attention to "how data is generative of new forms of power relations and politics at different and interconnected scales" (Ruppert et al., 2017, p. 2). These and other authors underline that inequality and power relations are not just effects of big tech operations but permeate many of the processes connected with the deployment of data and data infrastructures. Critical data studies suggest that data should not be viewed as merely a technical issue, heralding more efficiency, transparency, and analytical precision, but that data are intermeshed with complex organizational, social, political, and historical contexts and issues. Furthermore, they suggest that scholars must study *in action* how data is imagined, produced, contested, and displaced across a variety of fields and cases (boyd & Crawford, 2012; Bates et al., 2016; Dalton et al., 2016).

Following the calls to study the organizational and political aspects of recent investments around data, this study provides insights into the construction of a public data infrastructure and how this development is involved in changing public-private relationships. I refer to data

infrastructures as a specific kind of information infrastructure[1] with the primary purpose to share data across organizations. Several researchers have argued that when public organizations invest in data infrastructures and information technologies, they risk becoming locked into relationship with private suppliers, and that the boundaries between public and private organizations become blurred (Pasquale, 2015; Tréguer, 2019; Löfgren & Webster, 2020). Based on a mapping of Government discourses on big data in the Canadian Government, Joanna Redden finds that public-private partnerships are rapidly changing but are an area of silence that needs greater attention from academics and citizens (Redden, 2018). As will become clear throughout the analysis, it is no easy task to untangle the many ways public and private actors are enmeshed in the development of data infrastructures. Following Gray et al. (2018) notion of data infrastructure literacy, this paper seeks to participate in cultivating a sensibility for the social and political study of data infrastructures, and to study data infrastructures as relational "looking at how data infrastructures materially organise and instantiate relations between people, things, perspectives and technologies." (Gray et al. 2018).

Existing research have mainly focused on the use of data infrastructures, which have been studied in a wide range of settings including research (Bates, 2017; Plantin et al. 2018), health care (Wadmann & Hoeyer, 2018; Aula, 2019), education (Hartong & Förschler, 2019; Ratner & Gad, 2019), and social media (Bates, 2017; Gray et al. 2018). A common theme in several of these studies is the complications involved in making data 'flow', a difficulty that has been conceptualized as data friction (Edwards, 2010; Bates, 2017). Paul Edwards defines the term as "the costs in time, energy, and attention required simply to collect, check, store, move, receive, and access data" (Edwards, 2010, p. 84). Data friction refers to all that which hinders a 'seamless' exchange of data (Wadmann & Hoeyer, 2018). Most research into data friction has focused on the challenges of sharing and circulating data among diverse groups of actors; this touches, for example, on issues raised by interpreting meta data (Edwards et al., 2011), and the incentives of researchers to share their research data (Bates, 2017). As this article draws attention to data infrastructure in the making, it studies organizational and political frictions involved in the sharing of data, following Paul Edwards' hint that data frictions also have physical and social aspects and are just as much implicated by "the policies and practices of those who hold data and those who control access to data" (Edwards, 2010, p. 85).

This article investigates the Danish Customs and Tax Administration (DCTA) aspirations to take control over and responsibility for its data infrastructures. It is based on the ethnographic study

[1] Studies of data infrastructure often overlap with previous infrastructure studies (Star & Ruhleder, 1996; Star, 1999), and shares research objectives with studies of information infrastructures (Bowker et al. 2009) and knowledge infrastructures (Karasti et al., 2016).

of an IT project, commencing in 2016, to develop a new data-sharing infrastructure (DSI). I was inspired by discussions among my interlocutors who took part in the project to address the need to 'bring home' data and IT competencies. These discussions offered disconcertment (Verran, 1999; Law & Lin, 2010) in the sense that 'bringing home' at first seemed to indicate a clear shift in the control and responsibility of public data infrastructures from private to public actors. Through the analysis it became clear that 'bringing home' did not refer to any clear shifts as it was translated and discussed by public employees in different stages of the project. Before presenting the analysis, I offer some background information on why the DCTA wants to take back control over and responsibility for its data infrastructures. I then introduce my research approach and the analytical concept of reconfiguration (Suchman, 2007, 2012), which I have used to analyze my empirical material.

Organizational Background

> The plan that the government now presents constitutes enormous – yes, actually historically large – investments in the Danish Customs and Tax Administration. Whereas recent years have been marked by cutting expenditures, we now travel in the opposite direction. In the years to come, we will invest around 7 billion Danish Kroner in the tax administration, which will be used for more employees and new IT systems. In this regard, it is important to highlight that we have learned from the mistakes of our past. The failure of the EFI process should not make us afraid of the digital pathway. Digitalization is an inevitable part of the future. (Danish Ministry of Taxation, 2016[2])

The DCTA has been in a perpetual state of crisis for the past decade. Scandals have continued to pile up. Cases of tax evasion and outright fraught, as well as delayed and failed IT projects, have gathered widespread public attention in Denmark (Christensen & Mortensen, 2018). One of the most prominent examples is the failure of the IT system EFI. The development of EFI commenced around 2004 with the goal of automating the collection of debts owed to the Danish state. The following year, in 2005, the DCTA announced considerable restructuring and downsizing of the tax administration premised on the successful implementation of new IT systems, including EFI (Danish Ministry of Taxation, 2004). Between 2005 and 2015, the tax administration reduced the number of its employees from 11,500 to 6,500, a reduction of more than 40 percent (Herschend, 2015).

[2] Translated from Danish. Throughout the article I have translated the empirical data, such as reports and interviews, from Danish to English. My focus has been to keep the overall meaning rather than to translate the source verbatim.

EFI did not work as planned, however. In the Danish media, it became known as the 7-9-13 system (Høberg, 2015), ironically referencing a Danish saying, similar to 'knock on wood', supposed to prevent bad things from happening: 7, 9 and 13 indicate the number of years the project was postponed. When the system was finally put to use in 2013, it quickly became apparent that it was collecting a large number of debts that had expired and, consequently, that the system was operating illegally. In 2015, it was closed down. The total cost of the system at this point had accrued to more than one billion Danish Kroner (Mølsted, 2016); meanwhile, debts owed to the state continued to increase as they were not being collected.

In 2015, the crisis in the tax administration made the then Danish Minister of Taxation announce another major reorganization. Following a report from 2015, aptly titled *Tax administration away from the crisis* (Danish Ministry of Taxation, 2015), the organization was split into seven independent agencies (Danish Ministry of Taxation, 2016). One of these was the IT and Development Agency, employing approximately 1,000 employees across 80 offices, and one of the largest agencies in the new tax administration. In the report, despite their troubled past, digital technologies were yet again seen as crucial for responding to the ongoing crisis. Rather than retreating from the 'digital pathway', the tax administration instead chose to follow it further. This time around, it was emphasized that the DCTA wanted to take more control over and responsibility for the development and maintenance of its IT systems. The turbulent background of the DCTA, and its recent attempts to invest in data infrastructures and IT competencies, forms the backdrop of this discussion.

Research approach

As mentioned, this article is based on ethnographic fieldwork for the DSI project, which I carried out between November 2017 and December 2018. The methodological approach employed relied on the dictum from actor-network theory, that of follow[ing] the actors, which consists in paying detailed attention to events and happenings in the field (Latour, 2005). As noted by Winthereik (2020) this dictum "does not outline any specific kind of action, nor does it explain anything. It is an invitation for in situ sense-making and sorting out relations and attachments" (Winthereik, 2020, p. 24). I use it here in two senses: first, it describes the ethnographic approach of 'shadowing' project participants (Czarniawska, 2004); second, it refers to being inspired by the worlds and ideas of the project participants and making these a central part of the analysis. As Latour writes, it is the purpose of the analyst to 'catch up' with the 'wild innovations' of informants "in order to learn from them what the collective existence has become in their hands, which methods they have elaborated to make it fit together, which accounts could best define the new associations that they have been forced to establish" (Latour, 2005: 12).

The analysis that follows is based on participant observation in 34 project meetings and informal conversations with the DSI project participants over lunch, colloquially in the hallways, and between project meetings. During the period of data collection, the project went through the initial phases: the analysis phase, which concluded with the approval of a project initiation document in February 2018; and the acquisition phase, which ended with the completion of a tender process and the hiring of a group of external developers in June 2018.[3] The article also draws on ten semi-structured interviews conducted with the project's three main customers and seven project participants from the DCTA. Finally, I draw on central project documents such as the project initiation document, tender materials, and PowerPoints.

I approach this ethnographic material through the analytical concept of reconfiguration (Suchman, 2007, 2012), using it to analyze how boundaries between private and public actors shift and are negotiated discursively as well as materially. According to Lucy Suchman, configuration is "a tool to think with about the work of drawing the boundaries that reflexively delineate technological objects, and as a conceptual frame for recovering the heterogeneous relations that technologies fold together" (Suchman, 2012, p. 48). The notion of configuration is made up of two words: configure and figuration. To configure something refers to the act of arranging and ordering its elements in a particular form for specific purposes. Suchman main focus is how the boundaries between humans and machines and reconfigured in particular technologies and visions (Suchman, 2007). With the notion of figuration, Suchman highlights that arranging and ordering something does not just involve its physical parts, it also involves changes in related language and signs that set expectations, goals, and motivations. Reconfiguration is thus a material-semiotic process (Haraway, 1988). As we will see, the reconfiguration of data infrastructures is inextricably linked to the drawing of new boundaries between public and private actors in the DCTA. The notion of 'bringing home' is the central figuration that I explore in this article.

Analysis

The figuration of bringing home

In February 2017, The DCTA issued a job advertisement searching for an office manager for a new IT department; it was worded as follows: "The Danish Customs and Tax Administration are in the midst of building a new internal IT development organization in order to bring home a

[3] In project management parlance, these phases are referred to as 'the decision phase': "In the decision phase ministries will go through several activities: deciding on a project's goals, estimating a timeline and a budget, doing market analysis, choosing a development method, formulating requirements, and completing a tender process." (National Audit Office of Denmark, 2020: 3)

series of business-critical IT systems" (Stensdal, 2017). The job advertisement marked a new IT strategic direction for the DCTA, as the deputy director of a new IT department in the DCTA elaborated in an interview, "First of all, it is an alternative to the procurement and outsourcing paradigm, which we have had for better or worse." (Stensdal, 2017).

The announcement that the DCTA would bring home its IT systems was criticized by a number of associations in the Danish tech industry that did not support the 'creation of public monopolies', and opposed the public sector 'stealing tasks' from private companies. A spokesperson for the organization DI Digital, an association of private tech companies in Denmark, commented on the new strategy by stating, "The public sector should focus on its core competencies. It should not try to build public IT monopolies. Experience shows that if the public sector develops its own IT solutions, we get more expensive and poorer solutions, and we miss out on opportunities for growth and export in the private sector" (Version 2, 2017).

In the media, the figuration of 'bringing home' business-critical IT systems represents a potential conflict between the public and private sectors. Using the notion of DCTA as a home, a clear boundary is drawn between 'public' on the one hand and 'private' on the other. The 'home' of public organizations becomes separated from its 'away', which is constituted by private organizations. Bringing home business-critical IT systems indicates that the public sector is able to, will, and maybe even should take control and ownership of the IT systems currently managed and maintained by private suppliers.

During fieldwork, it became clear that 'bringing home' also figured as a point of orientation for employees in the DCTA in the development of a new data-sharing infrastructure (DSI). However, attending to the work involved in planning and constructing this data infrastructure reveals some complexities in how boundaries between public and private actors are re-made in practice. The result of bringing home was far from a clear shift in responsibility and control.

In the following analysis I first introduce the DSI project being analyzed, then turn to the work and discussions of the public employees involved in planning and designing the DSI. I focus on two aspects of the development wherein the figuration of 'bringing home' came up: first, how the DSI would relate to the larger infrastructural environment in which it would be embedded; and second, the difficulties the public employees and project participants faced in their wish to rely on internal competencies to manage and develop the data infrastructure.

The DSI project: Taking control of data sharing

Since the 1980s, it has been common practice to outsource public IT systems and solutions to private companies (Margetts, 1999; Landsberger & Wolken, 2001; Bates, 2012). "As the government privatized public assets and encouraged the outsourcing of public services, datasets needed by public bodies became increasingly owned or managed by private interests that extracted profit by selling data back to public authorities or demanding payment to undertake data retrieval" (Bates, 2012). It was against this trend that the DCTA was now proposing to work.

In November 2017, the Danish IT council[4] and the DCTA's upper management approved the project initiation document (PID) for developing the new DSI. The project's main goal was to enable a cheaper and more flexible way of sharing tax data across the public sector as a whole. The data infrastructure was described as the "foundation for the tax administration's future data sharing practices" and as the "first step towards simplifying the tax administration's overall data sharing for the benefit of businesses and public organizations", as well as contributing "to societal growth and creating a modern public sector" (PID, 2017).

The sharing of tax data is a central feature of the Danish digital welfare state. Government payouts and welfare services – including pensions, unemployment benefits, and state education grants – are calculated based on the recipient's income. Organizations such as pension and unemployment funds and municipalities thus rely on tax data to calculate whether citizens and companies are eligible for public services. Sharing tax data across the public sector is not entirely new. Since the beginning of the 21st century, the DCTA had outsourced the sharing of tax data to the IT supplier KMD. KMD was established in 1972 as the Danish municipalities' primary supplier and was originally a public organization (Dunleavy et al., 2004, p. 133; Jæger & Pors, 2017, p. 154). When KMD was privatized in 2008, some employees in the DCTA started to question whether KMD should still be in charge of sharing tax data. One of the tax employees explained the situation, "It felt like we just delivered all our data to KMD, which then further distributed it… It made a lot of sense in the old days, back when KMD was a public organization, but then they suddenly became a private organization" (Interview, 3rd of July, 2018).

The DCTA commenced the DSI project for two main reasons. First of all, the current data-sharing arrangement was expensive for the organizations who received tax data. Employees in the DCTA highlighted that KMD had a profit motive, making them charge as much as possible

[4] The Danish IT Council provides support to the IT projects of the Danish state. IT projects with budgets that exceed 10 million DKK require IT Council approval and must provide a risk assessment (https://digst.dk/styring/statens-it-raad/raadets-opgaver/).

for sharing the data. If the tax administration took charge of operating their own data-sharing infrastructure, then they would only need to cover development and maintenance costs. The second reason was that, after the GDPR legislation, the DCTA could no longer ignore discussions on responsible data practices and data protection (McDermott, 2017; Hartman et al., 2020). In one interview, an employee from the DCTA emphasized that the GDPR had sparked discussions about whether it was safe and responsible for a public organization to outsource its data-sharing practices. The DCTA was not required to develop a new DSI, but employees thought it would signal responsibility and improve the reputation of what had become a controversial public organization. When two organizations – Payout Denmark and the Danish Municipalities' IT Association – proposed to fund a new DSI, the tax administration was ready to accept their offer.

The DSI project represented a clear shift in public data-sharing practices. Prompted by economic, legal, and moral rationales, the DCTA wanted to take back control over and responsibility for the sharing of tax data. The project's Product Owner elaborated on this shift when I asked her what was new about the DSI project:

> We take control, instead of handing it over to someone else. We will know who gets what data down to the smallest detail, and we control our customers through data-sharing agreements. These agreements are not a legal requirement, but they are an excellent way to register who uses our data and for what purposes. (Interview, 3rd of July, 2018)

Taking control of data-sharing practices mapped onto the notion of bringing home business-critical IT systems. With the development of a new DSI, the tax administration would institute data-sharing agreements, providing it with an overview of its data recipients. The development of the DSI thus promised to solve immediate and practical concerns and to shift control over data-sharing from private to public actors. Examining the work involved in planning and developing this data infrastructure complicated the idea that 'bringing home' designated a process of drawing clear and simple boundaries between public and private actors.

I now turn to the initial phases of the project.

The Databank: Bringing data home

A precondition for developing the tax administration's new DSI was that the project would have access to tax data. Since the DCTA had outsourced its largest IT systems to private companies, the project relied on establishing a new database environment: The Databank. The project

participants talked about the Databank project as a process of bringing data home, but it was not clear exactly how the relationship between the tax administration and its private suppliers would be reconfigured.

> The people in the room burst into laughter as the project manager told them about his experience last week. He had visited the legal department, where one of the legal employees of the DCTA had been confused about a sentence used in the project: "What do people mean when they talk about bringing data home? How on earth is it possible to break contracts with the existing suppliers who manage and maintain the tax administration's IT systems?" The laughter in the room indicated that the legal employee had missed something obvious. The project manager made it clear that bringing data home did not mean that the tax administration would change the underlying IT systems. Instead, they would copy data from those systems, which would provide the basis for sharing data with external organizations. One of the project participants chimed in and added that maybe it would be better to talk of replicating data rather than bringing data home — a suggestion that brought about another round of laughter. (Observation, 22th of July, 2017)

According to STS scholar Helen Verran, laughter can indicate an interesting ethnographic moment – an experience where something feels amiss (Verran, 1999). To me, the laughter and the following exchange of words highlighted the confusing relation between the language used by employees in the DCTA and the ongoing configuration of new organizational relations and data infrastructures. The figuration of bringing data home apparently confused the lawyer, who seemed to interpret it to mean that the DCTA was looking to become independent from its IT suppliers. When another tax employee suggested speaking of replicating data rather than bringing data home, she highlighted that the imaginary was not one of data being removed from external suppliers, but of being replicated into a database environment under the control of the tax administration. The notion of replicating data clearly communicated that the tax administration would still be dependent on its external suppliers rather than breaking with them entirely. One of the project participants further elaborated:

> DXC administers and maintains the IT system for yearly income assessments – and it will continue that way. KMD administers and maintains the IT systems for the preliminary income assessment – and it will continue that way… These are the two suppliers that maintain some of the old IT systems. Our philosophy is, well we have called it 'bringing data home', but we will be

> replicating data [to store] in our databank in order for us to have them ready when someone asks for them. Then we don't have to get the data from KMD and DXC and pay them every single time we make a request. We can do something ourselves and be more directly in control, which will also make it cheaper. So that has been our philosophy. (Interview, 3rd of July, 2018)

The development of the DSI depended on a larger ecology of infrastructural arrangements in the tax administration. The DSI would be embedded in this ecology and be "built on an already installed base" (Star, 1999). This fact points to some of the complex organizational changes involved in the tax administration's efforts to take control of data infrastructures. Before the development of the DSI, KMD administered both the underlying IT systems and the data-sharing infrastructure. With the development of the Databank and the DSI, the tax administration could be seen as placing itself between KMD and the organizations receiving tax data. As a result, future DCTA personnel might find themselves negotiating with its external suppliers about how data is exchanged, something previously performed by the organizations receiving tax data, such as the municipalities and Payout Denmark. This can be seen as increasing the organizational complexity involved in sharing tax data.

The DSI project was involved in redrawing the boundaries between the tax administration, its private IT suppliers, and its data recipients. Through this particular infrastructural configuration, the employees in the DCTA hoped to gain control over data exchanges without re-developing the underlying IT systems. The process of taking control of data-sharing practices and infrastructures is one of establishing and 'cutting relations' (Ratner & Gad, 2019), and of "distinguishing that which belongs from that which is outside" (Bloomfield & Vurdubakis, 1999). In this process, the DCTA views data as material which can be controlled separately from the underlying IT systems that produce it, a rationale that can be understood in relation to the data revolution and the proliferation of data talk, as I outlined in the introductory section. As long as data are 'brought home', data-producing IT systems can 'stay away', an arrangement that reproduces a degree of reliance on existing IT suppliers, and, as discussed in the following section, requires the tax administration to engage in a range of new relations with new private actors.

Organizational maturity: Bringing home IT competencies

IT consultants have become ever-present in the Danish public sector. The Ministry of the State of Denmark recently reported that the public sector had spent 900 million DKK on consultants over four months in 2019 (Danish Finance Committee, 2020). The organization that had relied the most on IT consultants, by a wide margin, was the DCTA. With a total expenditure of 340 million DKK, it had contributed over one-third of the total outgoings (Ranhauge, 2020). Visiting

the DCTA, it was highly visible that the organization relied on consultants; in several offices, the first rows of desks were reserved for them and, just as often, they worked at desks alongside public employees. One of the participants in the DSI project jokingly told me, "There are consultants everywhere. Deloitte[5] must be the biggest public employer in Denmark" (interview, 3rd of July, 2018). In many ways, the tax administration was relying on private consultants to keep the organization as small as possible. Now the project participants were trying to figure out what it would mean for a public sector IT project to take greater responsibility for managing and developing IT systems.

An explicitly stated goal of the DSI project was to build capacity among internal employees. Indeed, the project manager talked about bringing home "the development of IT systems":

> We have the strategy that we want to bring home the development of IT systems. We cannot just place an order with a supplier and expect something in return. We need to think that we are in charge of managing, implementing, and maintaining the solution – so we have to consider all the processes and roles in the project. (Interview, 7th of July, 2018)

Although the project's goal was to rely on tax administration employees for the management and development of IT systems, the notion of bringing home the development did not mean that the project would not include private actors. On several occasions, the DSI project participants questioned whether the tax administration was sufficiently organizationally mature to handle the different tasks involved in managing and developing a new DSI. For example, it was mentioned in the project initiation document that "maturity within the tax administration with regards to DevOps[6] and agile methods is currently limited" (PID, 2017, p. 46). At one project meeting, the project manager, expressed concern about the maturity of the tax administration yet again, to which one of the other project participants responded, "But isn't it exactly this maturity that we are buying?" (Fieldnote, 25th of January, 2018).

To address the 'limited maturity' with regard to the software frameworks of DevOps and agile methods, the project hired two consultants to help participants understand and describe the DevOps principles (Tender documents, 2018). The project had also hired a second group of

[5] Deloitte is one of the largest consultancy companies in the world with more than 300,000 employees worldwide.

[6] DevOps is a software development method that emphasizes the connection between developing software and operating and maintaining software. The DCTA has a DevOps office, which is responsible for providing tools and software to assist in managing and developing IT solutions.

consultants to assist it with preparing its agile methodology and describing the DSI in accordance with agile principles, and they played a significant role in the initial phases of the project. The DSI project had hired them because of their expertise in developing data-sharing infrastructures in the public sector.[7] When these consultants spoke to me about their work in the project, they often used terms such as 'we' and 'us', as if they were as much an equal part of the project as the public employees. Their assistance included discussing the technical aspects of data-sharing with public employees, taking part in interviewing customers, writing user stories, and constructing tender documents for hiring external developers.

Interestingly, when the project participants considered bringing home IT development and competencies, they did not consider the opportunity this offered to avoid using consultants altogether. Instead, they kept questioning how they could include them while developing internal resources. The goal was to have internal employees involved in the project, learning from and collaborating with the consultants. An example to illustrate this point was how the project participants imagined using external developers in developing the DSI, suggesting two ways the project could receive external help while retaining expertise and know-how within the tax administration: first, they wanted the external developers to work inside the tax administration and as part of a team managed by tax employees; and second, developers from within the tax administration should work alongside the consultants. As the development of the DSI proceeded, the project could include more internal developers and slowly reduce the involvement of external developers.

Just as bringing data home did not designate a clear break with private actors, neither did bringing home IT development and competencies. While the project participants often expressed the desire to rely on 'internal employees', they still needed consultants for several key tasks, such as describing the requirements of the DSI and developing it. Bringing home IT development did not mean getting rid of consultants; rather, it seemed to suggest new ways of working with private actors. As the project manager further noted, "We bring home external developers, but we set the guidelines" (Interview, 7th of July, 2018). Bringing home IT competencies was not only about relying on internal skills but just as much about bringing consultants and external developers to work within the 'home' of the tax administration.

Discussion

There is no clear boundary between public and private organizations but a myriad of shifting relations and associations (Mitchell 1991; Treguer, 2020). Critical data studies have discussed how

[7] The consultancy firm had been part of developing the Data Distributor (Datafordeleren) a central public data sharing infrastructure in Denmark.

the increased use of and investment in data has affected the interplay between state and market forces (Maguire & Winthereik, 2019; Brevini & Pasquale, 2020). Many discussions focus on the consequences of algorithms and data infrastructures in use, whereas this article focuses on data infrastructure in the making. In this discussion, I address three related claims about how public organizations' interest in data affects the relationship between public and private actors. The three claims are: 1) That private companies gain increased power in relation to public organizations, e.g., by getting access to Government data and taking over public services – in other words, that investments in data are an intensification of neoliberal policies. 2) That public organizations have become highly dependent on private companies to manage and maintain their IT-systems, making it almost impossible for public organizations to take back control and responsibility of IT systems. 3) That data technologies' deployment is involved in blurring the boundaries between public and private actors. I now discuss these questions in relation to the construction of data infrastructures in the public sector based on the DSI case presented in this paper.

The figure of 'bringing home' has become prominent in the Danish Government and the DCTA. The figure is a direct response to decades of outsourcing the responsibility and control of public IT solutions. It thus provides a vision that stands in contrast to the many stories of how private organizations are taking over Government services and data. Instead, the figure shows how the increased focus on data can also lead to public organizations wanting to be more in control of their data to more effectively utilize it and be responsible for handling citizens' data.

But while the figure signaled a relatively straightforward aspiration for public organizations to take back the responsibility and control of critical IT systems and data infrastructures, the present case illustrated several frictions, which public employees face in their endeavor to take control of IT solutions. These frictions address the second claim presented above that public organizations have become locked into private dependencies that they cannot undo. First, the DCTA experienced political frictions. When the tax administration announced that they would bring home critical IT systems, several IT associations responded, warning that the Government shouldn't 'steal' tasks from private companies and should focus on its core competencies. Second, the DSI project participants experienced technical frictions. The tax employees realized that bringing data home would not make them independent from the existing IT suppliers, making them consider if it would be more accurate to speak of the process as replicating data rather than bringing data home. Third, the public project participants experienced organizational frictions in managing and developing the new data infrastructure. While the project wanted to rely as much on internal competencies as possible, the project participants often questioned the maturity of the tax administration, which at several moments led them to hire consultants and external developers to assist in the management and development of a new data infrastructure.

On the one hand, these three frictions indicate that public organizations such as the DCTA are indeed heavily dependent on private companies and locked into arrangements that cannot easily be undone. The case illustrates that the DCTA is dependent on IT suppliers who manage their central legacy systems and that the organization lacks the skills and capacities to develop and manage new IT systems. But on the other hand, the case illustrated that it is indeed possible to make changes to a public organization's existing IT landscape. By developing a new data infrastructure, the DCTA takes on the responsibility and control of its data sharing. Tax employees become responsible for making new data-sharing agreements and getting an overview of when and what tax data are shared with other organizations.

Finally, I address to which extent it is meaningful to speak of a blurring of the boundaries between public and private actors in developing public data infrastructures. First, the separation of data infrastructures from the underlying IT systems that generate data illustrates a particular way in which the boundaries between public and private actors can be characterized as blurred. When data are replicated from one IT system to another, it sets up an ongoing relationship between different organizations and IT infrastructures. In the DSI project, the bringing home of data was involved in reproducing, even extending, the tax administration's reliance on its external suppliers. Second, how public and private actors worked alongside one another in the tax administration and the recent push to bring consultants and external software developers to work inside the tax administration is another way in which it makes sense to speak of blurred boundaries.

But I will argue here that in the case of developing new data infrastructures in the public sector, the blurred boundaries between public and private actors are also a consequence of the language and abstractions used to describe the ongoing changes. As I traced how the figuration of 'bringing home' was translated from an overall organizational strategy into a figure that oriented public employees' work, it became clear that the relationship between public and private actors is changing at so many different levels and layers. The distribution of ownership, responsibility, and control does not shift all at once from public to private actors. As Andrew Feenberg points out in his critical theory of technology, "Technologies can thus be analyzed as layered phenomena, reaching from the heights of full-blown ideology down to the details of technical design. At each level, further layers appear, reflecting different degrees of abstraction." (Feenberg, 2016, p. 658). While following the figure of "bringing home" highlighted some of these layers of negotiation, there was still countless additional layers where these boundaries became reconfigured, such as the wide range of new private software products introduced by public employees to help to

manage and develop the DSI (Douglas-Jones et al., 2018), and discussion about where to host and store the DSI upon development.

Conclusion

This paper illustrates how boundaries between public and private actors are negotiated in the development of a public sector data infrastructure. It highlights that the relationship between public and private actors are reconfigured through a material-semiotic process where new figures of change interact with what is technologically and organizationally feasible, and vice versa. Public employees in the DCTA use the figure of 'bringing home' to orient their work in the development of a new data infrastructure. At an overall discursive level, 'bringing home' signals an attempt to take back the responsibility for developing and maintaining public IT systems from private actors. The figure suggests that a rather smooth shift is possible, and that there exists a clear boundary between the 'home' of the public sector and its 'away' the private sector. Turning to the actual development phase of the DSI it was clear that 'bringing home' was translated and took on more concrete shapes in relation to the technical and organizational issues at hand. At the technical level, bringing data home came to describe a separation between the underlying IT systems producing data, and the data infrastructures for storing and re-distributing the data. At the organizational level, 'bringing home' IT competencies did not refer to a clear break with private actors but a specific way of bringing in consultants and external developers to work within the offices of the tax administration, working alongside the tax employees.

The paper contributes to recent debates on data infrastructures with an empirical case providing insights into developing data-sharing infrastructures in the public sector. I suggest that it is a crucial task for creating an organizational and political literacy on data infrastructures (Gray et al., 2018), that we move beyond smooth narratives that describe the control and responsibility of public data infrastructures as a matter of being either located at public or private actors.

References

Aula, V. (2019) 'Institutions, infrastructures, and data friction – Reforming secondary use of health data in Finland', Big Data & Society. doi: 10.1177/2053951719875980.

Bates, J. (2012) '"This is what modern deregulation looks like" : co-optation and contestation in the shaping of the UK's Open Government Data Initiative', The Journal of Community Informatics, 8(2). Available at: http://ci-journal.net/index.php/ciej/article/view/845 (accessed 7 August 2020).

Bates, J. (2014) 'The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom', Government Information Quarterly, 31(3), 388-395. doi: 10.1016/j.giq.2014.02.009.

Bates, J., Lin, Y.-W. and Goodale, P. (2016) 'Data journeys: Capturing the socio-material constitution of data objects and flows', Big Data & Society. doi: 10.1177/2053951716654502.

Bates, J. (2017) 'The politics of data friction', Journal of Documentation, Vol. 74, No. 2, pp. 412-429. doi: 10.1108/JD-05-2017-0080.

Benbouzid, B. (2019) 'To predict and to manage. Predictive policing in the United States', Big Data & Society. doi: 10.1177/2053951719861703.

Bloomfield, B. P. and Vurdubakis, T. (1999) 'The Outer Limits: Monsters, Actor Networks and the Writing of Displacement', Organization, 6(4), pp. 625–647. doi: 10.1177/135050849964004.

Bowker, G. C. et al. (2009) 'Toward information infrastructure studies: Ways of knowing in a networked environment.' In: *International handbook of internet research*. Dordrecht: Springer, pp. 97-117.

boyd, D. & Crawford, K. (2012) 'Critical Questions for Big Data', Information, Communication & Society, 15:5, 662-679, doi: 10.1080/1369118X.2012.678878.

Brevini, B. and Pasquale, F. (2020) 'Revisiting the Black Box Society by rethinking the political economy of big data', Big Data & Society. doi: 10.1177/2053951720935146.

Christensen, J. G. and Mortensen, P. B. (2018) *Overmod og afmagt: Historien om Det nye SKAT*. Copenhagen: Djøf Forlag.

Christin, A. (2017) 'Algorithms in practice: Comparing web journalism and criminal justice', Big Data & Society. doi: 10.1177/2053951717718855.

Czarniawska, B. (2004) 'On Time, Space, and Action Nets', Organization, 11(6), pp. 773–791. doi: 10.1177/1350508404047251.

Danish Finance Committee (2020) FIU Alm.del - endeligt svar på spørgsmål 32 Offentligt.

Danish Ministry of Taxation (2004) *Det nye skatteministerium*. Copenhagen: Skatteministeriet.

Danish Ministry of Taxation (2015) *Skat ud af krisen – Et oplæg til politisk samling*. Copenhagen: Skatteministeriet.

Danish Ministry of Taxation (2016) *Et nyt skattevæsen. Ny organisering, flere medarbejdere og velfungerende IT*. Copenhagen: Skatteministeriet.

Douglas-Jones, R. et al. (2018) A bestiary of digital monsters. In *Working Conference on Information Systems and Organizations* (pp. 177-190). Springer, Cham.

Edwards, P. N. (2010) A vast machine: Computer models, climate data, and the politics of global warming. Cambridge: The MIT Press.

Edwards, P. N. et al. (2011) 'Science friction: Data, metadata, and collaboration', Social Studies of Science, 41(5), pp. 667–690. doi: 10.1177/0306312711413314.

Feenberg, A. (2017) A critical theory of technology. In: Felt et al. (eds) *The Handbook of Science and Technology Studies*. Cambridge: The MIT Press, pp. 635-665

Fiore-Gartland, B., & Neff, G. (2015) 'Communication, mediation, and the expectations of data: Data valences across health and wellness communities', International Journal of Communication, 9(2015), pp. 1466–1484.

Graham, S. & Marvin, S. (2001) *Splintering Urbanism: Networked Infrastructures, Technological Mobilities and the Urban Condition*. New York, NY: Routledge.

Gray, J., Gerlitz, C. and Bounegru, L. (2018) 'Data infrastructure literacy', Big Data & Society. doi: 10.1177/2053951718786316.

Haraway, D. (1988) 'Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective', Feminist Studies 14, no. 3: 575-99. doi:10.2307/3178066.

Hartman, T. et al. (2020) 'Public perceptions of good data management: Findings from a UK-based survey', Big Data & Society. doi: 10.1177/2053951720935616.

Hartong, S. and Förschler, A. (2019) 'Opening the black box of data-based school monitoring: Data infrastructures, flows and practices in state education agencies', Big Data & Society. doi: 10.1177/2053951719853311.

Heaphy, L. (2019) 'Data ratcheting and data-driven organisational change in transport', Big Data & Society. doi: 10.1177/2053951719867359.

Herschend, S. (2015) GRAFIK Se Skats store mandefald. *DR*. Available from: https://www.dr.dk/nyheder/penge/grafik-se-skats-store-mandefald (accessed 19 February 2021).

Hockenhull, M. & Cohn, M. (2020) 'Hot air and corporate sociotechnical imaginaries: Performing and translating digital futures in the Danish tech scene', New media & society , 1-20. doi: 10.1177/1461444820929319.

Hong, S. et al. (2019) 'Big Data and government: Evidence of the role of Big Data for smart cities', Big Data & Society. doi: 10.1177/2053951719842543.

Høberg, J. (2015) Analyse Sådan blev Skats hundedyre it-system til 7-9-13-systemet. *FINANS*, Available from: https://finans.dk/finans/erhverv/ECE8006002/S%C3%A5dan-blev-Skats-it-system-til-700-mio.-kr.-til-7-9-13-systemet/?ctxref=ext (accessed 19 February 2021).

Hoeyer, K. (2019) 'Data as promise: Reconfiguring Danish public health through personalized medicine', Social Studies of Science, 49(4), pp. 531–555. doi: 10.1177/0306312719858697.

Jensen, D. (2008) KMD solgt til kapitalfond. *Computerworld.* Available from: https://www.computerworld.dk/art/49198/kmd-solgt-til-kapitalfond (accessed 19 February 2021).

Landsbergen Jr, D., & Wolken Jr, G. (2001) 'Realizing the promise: Government information systems and the fourth generation of information technology', Public administration review, 61(2), 206-220. doi: 10.1111/0033-3352.00023.

Latour, B. (2005) Reassembling the social: an introduction to actor-network-theory. Oxford: Oxford University Press.

Law, J. and Lin, W. (2010) 'Cultivating Disconcertment', The Sociological Review, 58(2_suppl), pp. 135–153. doi: 10.1111/j.1467-954X.2011.01966.x.

Löfgren, K. and Webster, C. W. R. (2020) 'The value of Big Data in government: The case of 'smart cities'', Big Data & Society. doi: 10.1177/2053951720912775.

Karasti, H, Millerand, F, Hine, C. (2016) Knowledge Infrastructures, Parts I-IV. Science and Technology Studies 29(1–4).

Kitchin, R. (2014) *The data revolution: big data, open data, data infrastructures & their consequences.* London: SAGE Publications Ltd. doi: 10.4135/9781473909472.

Kitchin, R. (2017) 'Thinking critically about and researching algorithms', Information, Communication & Society, 20:1, 14-29. doi: 10.1080/1369118X.2016.1154087

Kitchin R., & Lauriault, T. P. (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 (accessed 19 February 2021).

Maguire, J. & Winthereik, B. R. (2019) 'Digitalizing the State: Data Centres and the Power of Exchange', Ethnos, doi: 10.1080/00141844.2019.1660391

Margetts, H. (1999) *Information technology in government: Britain and America.* London: Routledge.

McDermott, Y. (2017) 'Conceptualising the right to data protection in an era of Big Data', Big Data & Society. doi: 10.1177/2053951716686994.

Mitchell, T. (1991) 'The limits of the state: Beyond statist approaches and their critcs'. The American Political Science Review, 85 (1): 77. doi: 10.2307/1962879

Mølsted, H. (2016) Samlet regning for EFI-skandale: Over én milliard kroner. *Version2*. Available from: https://www.version2.dk/artikel/samlet-regning-efi-skandale-over-en-milliard-kroner-602459 (accessed 19 February 2021).

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

Plantin, J.-C., Lagoze, C. and Edwards, P. N. (2018) 'Re-integrating scholarly infrastructure: The ambiguous role of data sharing platforms', Big Data & Society. doi: 10.1177/2053951718756683.

Ranhauge, C. (2020) Staten har købt it-konsulenter for 900 millioner kroner på fire måneder: Skatteministeriet er topscorer. DigiTech PRO. Available from: https://pro.ing.dk/digitech/artikel/staten-har-koebt-it-konsulenter-900-millioner-kroner-paa-fire-maaneder (accessed 19 February 2021).

Ratner, H. and Gad, C. (2019) 'Data warehousing organization: Infrastructural experimentation with educational governance', Organization, 26(4), pp. 537–552. doi: 10.1177/1350508418808233.

Ratner, H. and Ruppert, E. (2019) 'Producing and projecting data: Aesthetic practices of government data portals', Big Data & Society. doi: 10.1177/2053951719853316.

Redden, J. (2018) 'Democratic governance in an age of datafication: Lessons from mapping government discourses and practices', Big Data & Society. doi: 10.1177/2053951718809145.

Ruppert, E., Isin, E. and Bigo, D. (2017) 'Data politics', Big Data & Society. doi: 10.1177/2053951717717749.

Redden, J. (2018) 'Democratic governance in an age of datafication: Lessons from mapping government discourses and practices', Big Data & Society. doi: 10.1177/2053951718809145.

Stalph, F. (2020) 'Evolving data teams: Tensions between organisational structure and professional subculture', Big Data & Society. doi: 10.1177/2053951720919964.

Star, S. L. (1999) 'The Ethnography of Infrastructure', American Behavioral Scientist, 43(3), pp. 377–391. doi: 10.1177/00027649921955326.

Star, S. L., & Ruhleder, K. (1996) 'Steps toward an ecology of infrastructure: Design and access for large information spaces', Information systems research, 7(1), 111-134.

Stensdal K (2017) Skat satser stort på ny DevOps-afdeling - klar til opgør med traditionelle it-udbud. Computerworld, Available from: https://www.computerworld.dk/art/239205/skat-satser-stort-paa-ny-devops-afdeling-klar-til-opgoer-med-traditionelle-it-udbud (accessed 19 February 2021).

Suchman, L. (2007) *Human-Machine Reconfigurations*. New York: Cambridge University Press.

Suchman, L. (2012) Configuration. In: Lury & Wakeford (eds) *Inventive Methods – The happening of the social*. London: Routledge, pp. 48-61.

Tréguer, F. (2019) Seeing Like Big Tech: Security Assemblages, Technology, and the Future of State Bureaucracy. In: Bigo, D., Isin, E. & Ruppert, E. (eds) *Data Politics: Worlds, Subjects, Rights*. London: Routledge, pp. 145-164.

Verran, H. (1999) 'Staying true to the laughter in Nigerian classrooms', The Sociological Review, 47: 136-155. doi: 10.1111/j.1467-954X.1999.tb03486.x.

Version2 (2017). DI Digital om ny udviklingsafdeling i Skat: Det er offentligt opgavetyveri. Version2, Available from: https://www.version2.dk/artikel/di-digital-ny-udviklingsafdeling-skat-offentligt-opgavetyveri-1073508?utm_medium=feed&utm_source=version2.dk&utm_campaign=it_nyheder (accessed 19 February 2021).

Wadmann, S. and Hoeyer, K. (2018) 'Dangers of the digital fit: Rethinking seamlessness and social sustainability in data-intensive healthcare', Big Data & Society. doi: 10.1177/2053951717752964.

Winthereik, B. R. (2020). Is ANT's radical empiricism ethnographic?. In: Blok, A., Farías, I. & Roberts, C. (eds) *The Routledge Companion to Actor-Network Theory*. London: Routledge, pp. 24-33.