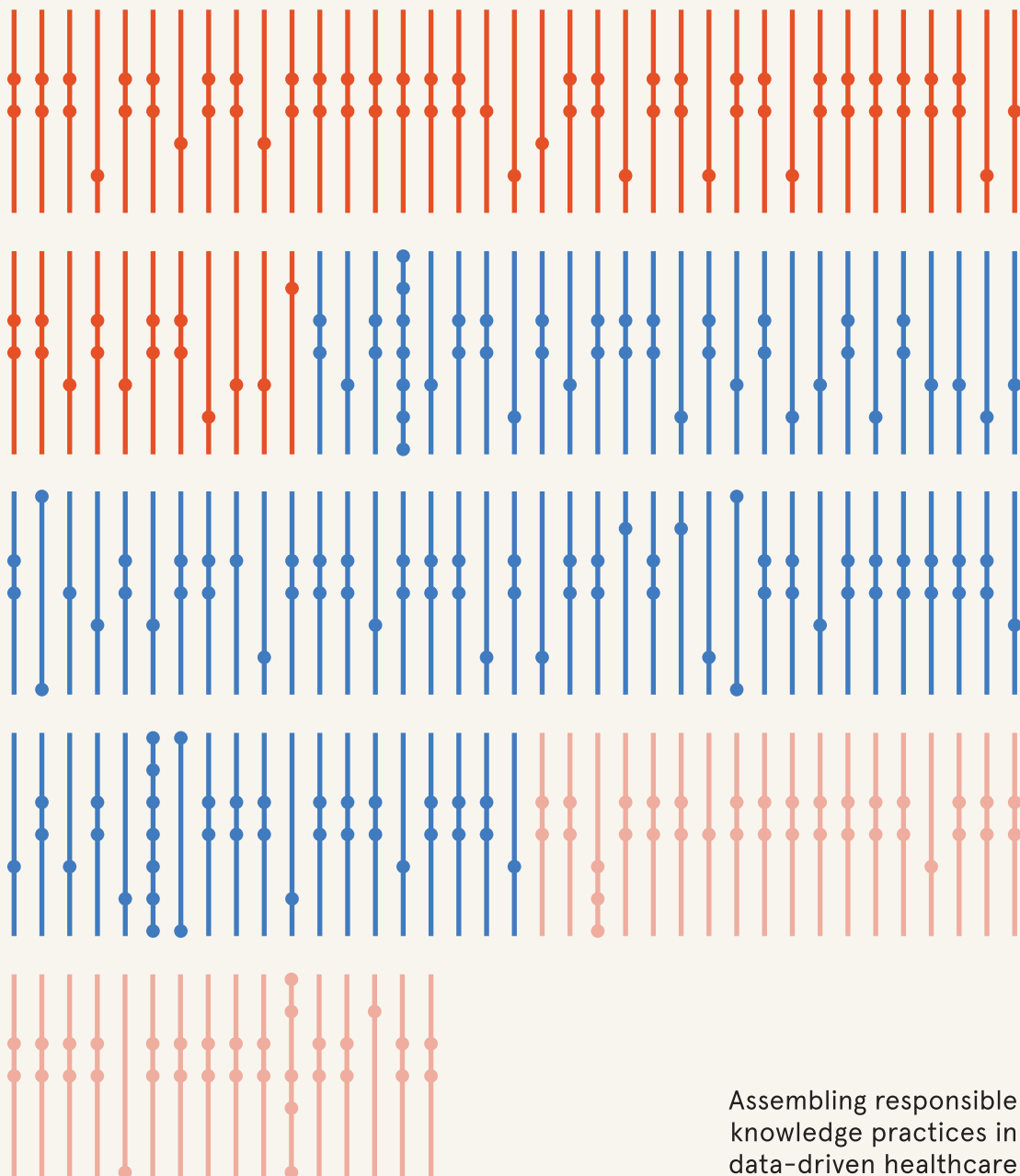


# Dreaming with data

Marthe Stevens



Assembling responsible  
knowledge practices in  
data-driven healthcare

Each vertical line on the cover represents 1 of the 164 people that actively contributed to 'Dreaming with Data'. The individuals are listed alphabetically and within their main area of expertise.

- Red lines are those interested in data.
- Blue are those interested in healthcare.
- Pink are those interested in society.

The 7 dots on the line indicate in which chapter each person played a role.

Marthe Stevens

# Dreaming with data

Assembling responsible  
knowledge practices in  
data-driven healthcare

The research for this dissertation was conducted at the Erasmus School of Health Policy & Management (ESHPM), Erasmus University Rotterdam. Chapters 2 and 3 in this dissertation were supported by the Horizon 2020 Innovation Program (grant number 780495; project BigMedilytics).

Copyright © 2021 by Marthe Stevens. All rights reserved  
ISBN: 978-94-6361-555-6

Cover design by Studio Suze Swarte, [www.suzeswarte.nl](http://www.suzeswarte.nl)  
Illustrations chapter 2 by Sue Doeksen, [www.suedoeksen.nl](http://www.suedoeksen.nl)  
Printing by Optima Grafische Communicatie, [www.ogc.nl](http://www.ogc.nl)

## **Dreaming with Data**

Assembling responsible knowledge practices in data-driven healthcare

## **Dromen met Data**

Bouwen aan verantwoorde kennispraktijken in data-gedreven gezondheidszorg

Thesis

to obtain the degree of Doctor from the  
Erasmus University Rotterdam  
by command of the  
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.  
The public defence shall be held on

Friday 4 June 2021 at 10.30 hrs  
by

Marthe Jacoba Stevens  
born in Haarlem, Netherlands.

## **Doctoral Committee**

Promotor:           prof.dr. A.A. de Bont

Other members:   prof.dr. R.A. Bal  
                          prof.dr. M. Huijer  
                          prof.dr. S.M.E. Wyatt

Copromotor:       dr. R.L.E. Wehrens

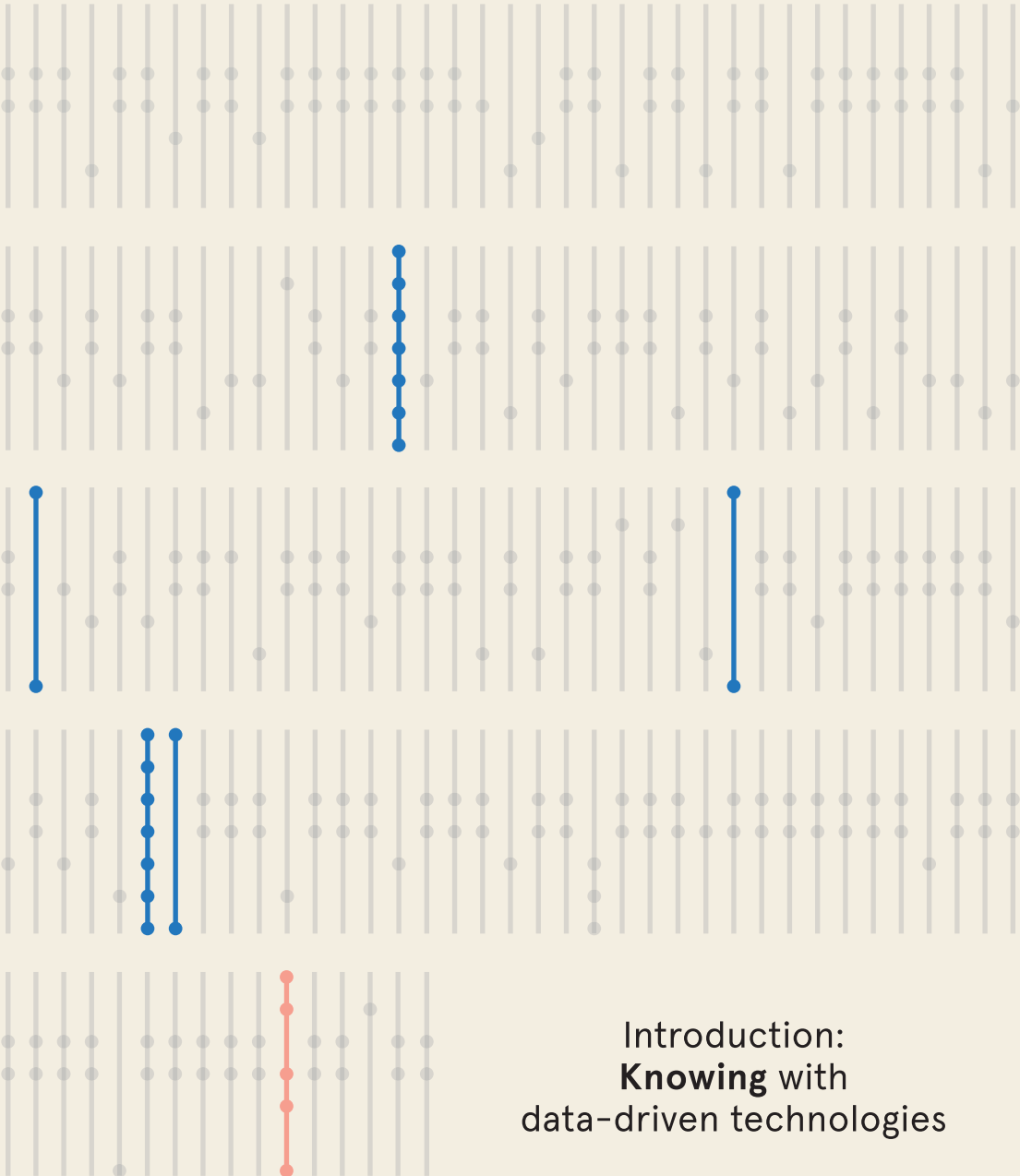
## Table of contents

Chapter 1	Introduction: <i>Knowing</i> with data-driven technologies	7
Chapter 2	<i>Perceiving</i> data-driven technologies	25
Chapter 3	<i>Framing</i> data-driven technologies	49
Chapter 4	<i>Dreaming</i> of data-driven technologies	69
Chapter 5	<i>Negotiating</i> on data-driven technologies	91
Chapter 6	<i>Caring</i> for data-driven technologies	113
Chapter 7	Conclusions: <i>Reflecting</i> on data-driven technologies	133
	References	157
	Summary	173
	Samenvatting	181
	Dankwoord	189
	Curriculum Vitae	197





# Chapter I



Introduction:  
**Knowing** with  
data-driven technologies



## Introduction

*"The era of Big Data challenges the way we live and interact with the world. Most strikingly, society will need to shed some of its obsession for causality in exchange for simple correlations: not knowing why but only what. These overturns centuries of established practices and challenges our most basic understanding of how to make decisions and comprehend reality. [...] Just as the telescope enabled us to comprehend the universe and the microscope allowed us to understand germs, the new techniques for collecting and analyzing huge bodies of data will help us make sense of our world in ways we are just starting to appreciate" (Mayer-Schönberger and Cukier, 2014: 6-7).*

There are high hopes for Big Data in healthcare. Innovative data analytics could have the potential to radically transform the field. Just as telescopes and microscopes opened our eyes to the universe and micro-organisms, data analytics are expected to enable us to reach new dimensions of reality and uncover new "truths" (Mayer-Schönberger and Cukier, 2014). According to some of the more vocal Big Data initiators, healthcare has been dominated by gut feeling and intuition for too long. Now it is time to be guided by data and rigor; to let "the data speak" instead (Mayer-Schönberger and Cukier, 2014: 6) and establish a new culture of decision-making (McAfee and Brynjolfsson, 2012). Inspired by these hopes, people dream that we can discover new associations, see new patterns and use these to make more personalized predictions, smarter decisions and more effective interventions. As a result, we would improve quality of care, save lives and lower healthcare costs (Raghupathi and Raghupathi, 2014).

These kinds of dreams have led to a surge of interest in data analytics for medical decision-making. Technically it is possible to collect more data from various sources and better analytical methods are available to meaningfully process these data (Raghupathi and Raghupathi, 2014; Kruse et al., 2016). Inspired by commercial successes, healthcare initiatives are starting to use data to detect diseases at earlier stages (Raghupathi and Raghupathi, 2014), predict the next virus outbreak (Mayer-Schönberger and Cukier, 2014) and tailor medical treatments to the needs of individual patients (Raghupathi and Raghupathi, 2014).

But not everyone shares these dreams for data in healthcare. Take, for example, the following quote:

*"Big Data creates a radical shift in how we think about research [...]. [It offers] a profound change at the levels of epistemology and ethics. Big Data reframes key questions about the constitution of knowledge, the process of research, how we*

*should engage with information, and the nature and the categorization of reality [...]. [There is] an arrogant undercurrent in many Big Data debates where other forms of analysis are too easily sidelined. Other methods for ascertaining why people do things, write things, or make things are lost in the sheer volume of numbers. This is not a space for welcoming the older forms of intellectual craft" (boyd and Crawford, 2012: 665-666).*

This quote emphasizes some of the dangers of a "radical shift" toward Big Data that are seen. The critics point out that data promises come with the risk of sidelining many established scientific methodologies, whereas many methodological issues remain, in their opinion, relevant to Big Data. Moreover, critics fear that innovative data analytics lead us to see patterns in data where none exist, that limitations and biases in data will be amplified when we cannot oversee the process (boyd and Crawford, 2012; Morley et al., 2020), leading to incomplete, inscrutable and misguided decision-making by medical professionals and healthcare policymakers with possible detrimental consequences (Beer and Grote, 2019; Househ et al., 2017).

The fears find their way into the scientific literature and popular culture. Publications describe how data analytics will reinforce discrimination and enlarge social inequality (O'Neil, 2016), how painful, wrong conclusions are drawn from limited healthcare data (Ebeling, 2016) and how using data analytics could challenge the authority of medical professionals and promote defensive decision-making which might harm patients (Beer and Grote, 2019). These fears are fueled by repetitions of all sorts of modern myths, on how misusing out-of-control technology is leading to harmful effects. Critics often refer to books on dystopian futures, when privacy is gone and technology has taken over control, such as *Brave New World* (Huxley, 1931), 1984 (Orwell, 1949) and movies such as *Minority Report* (Spielberg, 2002) (e.g. Gallagher, 2013; Mor, 2014; Robertson and Travaglia, 2019).

What Big Data optimists and pessimists have in common is that both perceive changes in the way we obtain knowledge in healthcare. Many proponents imply that over time, data-driven initiatives will lead to fundamentally new, improved ways of information gathering and decision-making. Meanwhile, many critics touch upon epistemological concerns, such as bias, lack of transparency and misguided information that leads to potentially dangerous decisions.

This dissertation argues that it is unproductive to approach Big Data in terms of hopes and fears in anticipating and shaping the future of healthcare with data-driven technology. Earlier studies on technological innovations have shown us that sticking to dualisms is unproductive as it obscures the more nuanced and subtle shifts underway (e.g. Ames,

2018; Smits, 2002). We know from research in fields such as Science and Technology Studies (STS) and philosophy of technology that the effects of large technological innovations will likely be *both* positive and negative in often highly situational ways – and probably will be different than expected (e.g. Janssen, 2016; Smits, 2002; Van Lente, 2012). This research project set off to produce a nuanced picture of Big Data by reflecting on changes in knowledge production in healthcare.

## Fleeting, nebulous technology

The idea to re-use and reconnect healthcare data gained momentum under the term “Big Data” around 2014. Many scholars tried and failed to define the term. Most famous was “the 3V” definition that characterized Big Data by its volume, velocity and variety (Mayer-Schönberger and Cukier, 2014). Big Data would consist of enormous quantities of data, created in real-time and could even be processed if it varied in format and quality. This was different from the data already being used in healthcare. That was smaller in quantity, time-consuming to processes and often produced in more controlled ways (Mayer-Schönberger and Cukier, 2014; Kitchin and McArdle, 2016).

Believing that the “3V” definition was too vague, many scholars attributed other qualities to Big Data to better describe what was going on. They proposed various extensions with extra “V’s”, such as value (many insights can be extracted and data can be repurposed) and veracity (data can be messy, noisy and contain uncertainty and error) (Kitchin and McArdle, 2016; Marr, 2014). Others argued that “V words” should be let go of, once and for all. They should be replaced by qualities such as exhaustivity (an entire system is being captured) and extensionality (can easily be extended) (boyd and Crawford, 2012; Mayer-Schönberger and Cukier, 2014). However, none of these descriptions seemed to last and what remained was a striking lack of ontological clarity about the term Big Data. The term acted like an “amorphous, catch-all label for a wide selection of data” (Kitchin and McArdle, 2016: 1).

Although the term Big Data has not lasted long, ideas about using innovative new data technologies for assembling and organizing data in healthcare have stayed. Now there are many terms to describe the trend toward intensified data use in healthcare. Concepts such as algorithms, artificial intelligence, data science, deep learning, (supervised and unsupervised) machine learning and predictive analytics are used almost interchangeably in many managerial and policy discourses (Rieder, 2018; Mayer-Schönberger and Cukier, 2014; McAfee and Brynjolfsson, 2012). Strictly speaking, all these terms refer to a range of computational methods and techniques to process data. But the underlying narrative

remains remarkably similar; the promise of using healthcare data innovatively with the help of computing-intensive statistical techniques to gain more detailed, complete and timely information that can be used for healthcare prevention and personalization (Kruse et al., 2016; Raghupathi and Raghupathi, 2014).

The object of study in this dissertation has shifted from the “strictly” defined, hyped techniques such as Big Data to focus on data-driven healthcare. The dissertation uses “data-driven technologies” to describe the general trend to data-driven healthcare. This approach to studying the techniques fits in with Critical Data Studies (CDS), Critical Algorithm Studies (CAS) and STS (e.g. boyd and Crawford, 2012; Wyatt, 2016). Scholars in these fields argue for the importance of studying Big Data as a broad cultural, technological and scholarly phenomenon that includes technological, analytical and mythological dimensions (boyd and Crawford, 2012). Translated to this dissertation, the term data-driven technologies refers to three things at once: to *new technological capabilities* to bring more, diverse sorts of data together in aggregated datasets; to the ability to conduct *new sorts of analyses* of datasets and use the results to make all sorts of claims in healthcare; and to the *mythology* surrounding the developments. For example, that modeling reality through data will produce better knowledge practices as it obtains insights blessed with “the aura of truth, objectivity and accuracy” (boyd and Crawford 2012: 663) or it will worsen practices as data it will be out-of-control and misused, leading to a variety of harmful effects.

Some people might be surprised by the explicit attention given to the mythology of data-driven technology. It is well known that technologies gain meaning and do particular things once they become embedded in practices. However, scholars in STS and the philosophy of technology have taught us that expectations, our hopes and fears for the future *also* profoundly influence the development of technologies and therefore require analytical attention. Visions of the future are often specific to time and place but occasionally grow into widespread imaginaries or ‘dreamscapes’ that seem to travel easily from context to context (Konrad et al., 2017; Jasanoff and Kim, 2015). These expectations, our dreams and imaginaries generate matters of concern, hide all sorts of normativities and acts as drivers for change, for example, by bringing certain actors together and steering research and investment agendas (Bensaude Vincent, 2014; Konrad et al., 2017). Thus the way we envision the data-driven future of healthcare has an influence on the choices we make today. This warrants critical scrutiny not only of the data-driven practices themselves, but also of the dreams, expectations and hopes that are attached to the technologies in various discourses.

## Knowledge practices in healthcare are being questioned

This dissertation foregrounds knowledge practices in healthcare as data-driven technologies question existing ways of decision-making, producing evidence and making sense of the world. Epistemology is the branch of philosophy concerned with the formation of knowledge. It studies the nature of knowledge, the rationality behind certain beliefs and asks such questions as: how do we know what we know? What does it mean to say that we say we know something? Why and how is knowledge accepted?

The healthcare sector is, in general, often characterized by a strongly institutionalized set of epistemological principles and accepted methodologies. The practice of knowledge production is often linked to the epistemology of evidence-based medicine (EBM), described as “the golden standard” for knowing and reasoning in medicine (Timmermans and Berg, 2003). The EBM tradition strives for objective, unbiased decision-making based on scientific methodology (preferably randomized controlled trials) and guidelines formulated for clinical decision-making based on the best available evidence (Van Baalen, 2019). In the last 30 years, these ideas led to many medical guidelines, a broad emphasis on accountability, transparency, standardization and control in healthcare, and improvements in healthcare quality and safety (RVS, 2017).

However, literature also points out that many knowledge practices in healthcare display much local variation (Bal, 2017). One important reason is that EBM relies on information produced in standardized situations; therefore it is detached from the daily practice of healthcare provision. For example, randomized controlled trials are often conducted under ideal, rigorously monitored conditions, which makes them only partly applicable to real-world settings. Because of this detachment, it is difficult to translate the generic knowledge produced by the EBM tradition to the diversity of individual patients, their personal values, and the particular setting of their care (Nicolini et al., 2007; RVS, 2017; Van Baalen, 2019). Thus, when making a medical decision, healthcare professionals often balance different sources, values and knowledge and generally prefer “personal”, “situated” and “local” knowledge based on their own or their colleagues’ experience, above the abstract detached information provided in scientific studies (Nicolini et al., 2007).

Another way of approach this is by focusing on the networks through which knowledge is formed. It then becomes visible that medical decisions are made in highly complex, entangled environments, in which various actors often collaborate across various disciplines. For example, a medical professional often works with technicians to evaluate a CT scan and develop knowledge about a patient (Nicolini et al., 2007; Van Baalen, 2019). This perspective also highlights that part of these networks are material and technological

objects that mediate our ways of knowing (Van Baalen, 2019; cf. Verbeek, 2015). Medical professionals gain knowledge – they know – because they use scientific guidelines, measure temperature with a thermometer and see malignant cells under a microscope.

The above shows that there is no one way of obtaining medical knowledge. There are many differences in knowledge practices, which have emerged over time, as medical specialties developed around particular scientific methodologies, diagnostics and interventions, and knowledge networks have formed. For example, in psychiatry, there is a great reliance on questionnaires used in combination with patients' narratives to characterize and communicate the patients' conditions (Ruppel and Voigt, 2019). In radiology, there is more attention for imaging techniques and the visual aspects of knowledge generation. Internal medicine pays more attention to biomedical measurements, clinical tests and blood measurements. These differences affect how knowledge is acquired and lead to differences in what is accepted as evidence.

Following their introduction, data-driven technologies have become part of the epistemic discussions going on in the fields of the healthcare sector. Many proponents seem to build on the assumption that data-driven technology will produce relevant information that adds value to healthcare practices. In their eyes, many current healthcare practices still "suffer" from too much uncertainty and unpredictability. New technological affordances make it possible to measure more aspects of our social world and turn them into data that are perceived to be "objective" and "true" (Crawford et al., 2014; Mayer-Schönberger and Cukier, 2014). These data can, for example, be used to tailor treatment to individual patients more precisely compared to the generalized information coming from medical guidelines and standard randomized controlled trials; thereby bridging gaps between science and practice. Moreover, the hope is that if we have enough real-world data, "the numbers can speak for themselves" (Anderson, 2008: 1) and facilitate a science driven by induction and reduction, without the need for theory and hypotheses (Kitchin, 2014; Mittelstadt and Floridi, 2016). By implication, we would no longer have to understand *why* certain things happen, but that we can measure or even predict *that* something will happen, thereby providing timely information even about phenomena that are impossible to study with current scientific methodologies.

However, critics fiercely resist this "reborn empiricism" (Kitchin, 2014: 3) as they fear that data-driven technology will produce problematic information that does not fit the ways knowledge is produced in healthcare. The critics argue that the new methodologies are not rigorous enough because healthcare actors may start to see and act on patterns in data where none exist. The outcomes of data-driven technology are probabilistic, not infallible; in contrast to randomized controlled trials they do not posit the existence of causal



relationships (Morley et al., 2020). In addition, critics worry that current limitations and biases in data will be amplified (boyd and Crawford, 2012), leading to misguided decision-making by medical professionals and healthcare policymakers, with possibly detrimental consequences (Beer and Grote, 2019; Househ et al., 2017). Lastly, they describe the technologies as opaque, increasingly self-learning and as a “black box” (O’Neil, 2016; Ziewitz, 2017), signaling concerns about the limited ability to know how certain conclusions are obtained and have oversight over the process. The critics say that this inscrutability makes knowledge creation in healthcare networks even more complex, harder to oversee, and unsuited to the healthcare culture that values personal, situated and locally produced knowledge (Beer and Grote, 2019).

To summarize, data-driven technology has become a topic of discussion on knowledge production and decision-making in the healthcare sector. More is going on than “just” introducing extra data or a new method that can be used in addition to current epistemic practices. Instead, it is raising questions on the sort of evidence that is necessary to make medical decisions, the importance of having theories and knowing where information comes from and how it is obtained. This calls for research that studies *how* data-driven technologies reconfigure knowledge practices in healthcare.

## What is a responsible knowledge practice?

Discussions on the best ways to obtain knowledge in particular healthcare settings are closely tied to ideas about what is deemed *responsible* knowledge practice. Many scholars have written about responsibility in relation to data-driven technologies. Writings include, for example, “eight principles for responsible machine learning and artificial intelligence” (The Institute for Ethical AI & Machine Learning, 2020) for addressing the ethical issues that arise with data-driven technologies. In addition, many organizations are publishing regulations (e.g. the General Data Protection Regulation, GDPR) and developing strategies for Responsible Research and Innovation (RRI) with data-driven technologies. For example, one strategy notes the importance of involving people and civil society organizations in the development of such technologies (Simon, 2015). All this work aims to set conditions and distinguish “responsible” from “irresponsible” data-driven healthcare.

While this work does offer valuable lessons, my dissertation takes another perspective. This dissertation argues that there is much to be learned about responsible knowledge practices by studying the use of data-driven technologies in healthcare (cf. Wyatt et al., 2013). Rather than looking for solid sets of rules and guidelines, hard criteria and fundamental

responsibilities, I empirically study how responsible knowledge practices are produced over time both through expectations and by actors in concrete practices.

I look at responsible knowledge from the perspective of *practices* for two reasons: (1) it accounts for the interrelatedness of epistemology and ethics better, and (2) it considers the affectivities and normativities that already play a role. Let me briefly explain.

The first reason is because many current discussions on the permissibility of data-driven techniques seem to belong to ethics as they seem to frame epistemology as part of the field. For example, they distinguish between such ethical themes as informed consent, privacy, ownership, inequalities and epistemology in assessing the literature on Big Data (Mittelstadt and Floridi, 2016). Because of the overwhelming attention given to issues of informed consent and privacy, this creates the potential of relatively neglecting the epistemological dimensions (Mittelstadt and Floridi, 2016; cf. Sharon, 2020; Wehrens et al., 2019).

In addition, framing epistemology as part of the field of ethics ignores the fact that ethics and epistemology are intimately linked to each other in many ways (cf. Simon, 2015). Consider, for example, much of what we think we know and the information that we decide to use influences what we do (or what we believe we ought to do) in a given situation and vice versa. All sorts of norms and values determine when it is good or permissible to hold a certain belief as true (Daston and Galison, 2007). Jasanoff (2004) describes this co-production of knowledge and our norms and values as:

*"[...] ways in which we know and represent the world (both nature and society) are inseparable from the ways in which we choose to live in it. [...] Scientific knowledge [...] both embeds and is embedded in social practices, identities, norms, conventions, discourses, instruments and institutions" (Jasanoff, 2004: 2-3).*

These ideas highlight the importance of attending to knowledge practices as socially produced and influenced by norms and values. This means that it is impossible to make harsh distinctions between ethics and epistemology.

I can illustrate this with a highly personal story by Ebeling (2017). She wanted to know why companies were marketing baby products to her, years after a miscarriage. She discovered that her very private healthcare data had ended up in a database of "new parents" that data brokers and marketers had commodified and sold to advertisers that wanted to market products directly to consumers.

This story is usually used to push the privacy debate (it shows the importance of keeping certain details about our lives private) and informed consent issues (even if we consent to handing over personal information, often we cannot oversee how it gets used and comes back to us). These debates are valuable, but what I want to highlight is that this Ebeling's story is about epistemology and its interrelatedness with ethics. In this case, the data brokers and marketers *believe they know* that a baby was born, which gives them the idea that they *are doing right* by presenting Ebeling with coupons for baby products and advertisements for preschools. Similarly, many of us find the commercial practices in this case *unethical*, because we *know* that the information is sadly not true. In another context, having other information available, we might have valued the offers of special discounts and free products.

To summarize, my first reason enables me to move beyond the “artificial” distinctions between ethics and epistemologies and consider their interrelatedness. The second reason for studying responsible knowledge *practices* is that it accounts for the affectivities and normativities that already play a role in healthcare practice (cf. Zuiderent-Jerak, 2007).

Current ethical and epistemological debates seem to focus on: (1) agenda-setting that makes particular ethical and epistemological dilemmas visible, for example, by highlighting key concerns in a particular context (e.g. biomedicine) (Mittelstadt and Floridi, 2016; Mittelstadt et al., 2016; Mittelstadt, 2019; Morley et al., 2020). (2) Theorization that leads to new conceptualizations of data-driven technologies, for example, on how normative and epistemic tradeoffs are made in theory (Grote and Berens, 2020). (3) Problematicization that highlights the misalignments between data-driven technologies and ethical concepts and principles we have in place. For example, scholars argue that traditional notions of moral agency should be reformulated in the context of self-learning systems (Floridi and Sanders, 2004).

These studies offer many valuable insights, yet by developing normative ideas relatively separately from practices that construct data-driven healthcare, they potentially neglect the affectivities and normativities present there. Instead of adding normative complexity to healthcare practice, I want to study the ideas and norms that are already in place and use them as a starting point for reflection on responsible knowledge practices in data-driven healthcare (cf. Zuiderent-Jerak, 2007).

## Research aim and questions

The aim of this dissertation is to critically investigate how data-driven technologies reconfigure what are deemed responsible knowledge practices in the healthcare sector. Important to note is that this research aims to understand how actors working on data-driven healthcare shape responsible knowledge practices themselves. This leads to the following central research question: *How do various actors reconfigure responsible knowledge practices in data-driven healthcare?*

This aim is translated into three sub-questions:

1. *How do actors in data-driven healthcare envision responsible knowledge practices?*

Data-driven technologies are bound by many hopes and dreams, and fears and critiques that make all sorts of epistemic claims about the future of healthcare. From the literature, we know that these expectations for the future have all sorts of influence on way data-driven technologies conceived, managed and enacted today. This question seeks to investigate how responsible knowledge practices are envisioned and what the consequences are.

2. *How do actors in data-driven healthcare construct responsible knowledge practices?*

Data-driven technologies are increasingly used in initiatives that aim to capitalize on the promises of data-driven healthcare. In day-to-day work of actors within healthcare initiatives, technologies receive meaning and do particular things. This question seeks to provide insights into how responsible knowledge practices are constructed in such initiatives.

3. *How are the roles and responsibilities with regards to knowledge practices in data-driven healthcare reconfigured?*

Data-driven technologies are questioning the existing knowledge practices in healthcare. This final sub-question aims to provide an understanding of how knowledge practices in data-driven healthcare are shifting and changing the allied roles and responsibilities.

## Investigating knowledge production in practice

The field of STS has much experience with studying knowledge practices, which explains why this body of work is an important source of inspiration for this dissertation. STS scholars critiqued the idea of an “epistemic unity of the sciences” (Galison and Stump, 1995). In the 1920s and 1930s, when fascism arose and there was growing tension between states, scholars of the Vienna Circle (*Wiener Kreis*) promoted the notion of a united science. They argued that there was only one kind of knowledge, only one science and only one scientific method and hoped that “an international scientific worldview could curb the divisive racial

and nationalistic worldviews" (Galison, 1995: 6). Many contemporary scientists still support versions of the unity-of-science thesis.

Over the years, scholars argued against these ideas. Instead, they contended that scientific practices are a far broader terrain; there is no one way of doing science, but a diversity of scientific practices, in which various, often local norms and values play a role and guide what is perceived to be normal and acceptable (Galison and Stump, 1995; Kuhn, 1962; Latour, 1987; Pinch and Bijker, 1984).

In the 1970s and 1980s, many STS scholars began studying the concrete practices in which scientific knowledge is obtained. They used ethnographic fieldwork to conduct a range of 'laboratory studies' to investigate the mundane day-to-day interactions through which knowledge is constructed, scientific research is done, and "facts" are produced (Knorr Cetina, 1981; Latour and Woolgar, 1986). They made it their mission to think through the encounters of disparate knowledge traditions and developed various concepts to study the diversity of epistemic practices. Some of these concepts are particularly useful in interpreting and understanding the rest of this dissertation.

First, looking at epistemic practices Knorr Cetina (1981; 1999) developed the concept of "*epistemic cultures*". She is known for her ethnographic comparison between experimental high-energy physics and molecular biology and summarizes the differences in strategies and policies of knowing between both fields. According to Knorr Cetina, these fields or epistemic cultures are "amalgams of arrangements and mechanisms – bonded through affinity, necessity and historical coincidence – which, in a given field, make up how we know what we know" (Knorr Cetina, 1999: 1). The notion of culture helps to see knowledge construction that takes place in a concrete setting as something in relation to a certain tradition and affinity.

Second, epistemic norms and values guide what is perceived to be normal and acceptable in epistemic practices and cultures (Daston and Galison, 2007; Latour, 1987). Daston and Galison (2007) developed the concept of "*epistemic virtues*" as a way of highlighting these norms and values. Their work on the notion of "objectivity" studied the meaning of the epistemic virtue of "objectivity" and described how it was understood differently by scientists in the 18th, 19th and 20th centuries. They argue that the interpretation of this virtue changed together with the practices and cultures of doing science. As a result, what was perceived as a "responsible" or "good" scientific practice throughout these ages has also changed.

Third, epistemic cultures and practices are dynamic domains of social life that are not closed off from others. Galison foregrounded language in analyzing how distinct communities in physics – such as theorists, experimentalists and engineers – come together in “*trading zones*” and create in-between vocabularies that facilitate communication and alignment of activities (Galison, 1997). Intermediating languages can range from simple (interlanguages) to complex (“pidgin”) and eventually, a shared language can emerge (“creole”) (Collins et al., 2007; Galison, 1997). These languages make it possible to interact and exchange goods despite differences and without homogenizing the inherent diversity in their communities (Galison, 1997).

## The research trajectory of this dissertation

This PhD research project began with a focus on Big Data, but Big Data proved to be nebulous. When I started the project in 2016, some people in the healthcare field told me that “Big Data was already out of fashion” and that now “they only spoke of machine learning.” I quickly realized that the shift toward data-driven healthcare should become the object of study in this dissertation and not “strictly” defined, temporarily hyped techniques such as Big Data.

This move was necessary as I noticed that different vocabularies led to boundaries between disciplines, actors and organizations. It is well known that practices of defining are largely about demarcating those who can play a role or have a say in certain developments from those who cannot (e.g. Seaver, 2017). I also saw this happening: “Big Data” conferences were deemed unattractive for “machine learning” experts even if similar themes were discussed. Being flexible in the inclusion of the various technologies meant that I could empirically explore data-driven healthcare practices *across* different disciplines, actors and organizational boundaries.

In line with the STS tradition of studying knowledge practices, I based this research project on an ethnographic sensibility. However, I realized that capturing the reconfiguration of knowledge practices through various nebulous, fleeting technologies required multi-sited ethnography (e.g. Hine, 2007; Marcus, 1995). This is the most suitable approach when an object of study is not bound to one site. It enabled me to follow an object as it circulates through institutional sites and epistemic practices.

Many authors in CDS and CAS argue that fleeting and nebulous data-driven technologies create particular methodological challenges for empirical research (Burrows and Savage, 2014). They argue that data-driven methodologies are often portrayed as powerful yet

inscrutable entities that govern, shape and control our lives in unprecedented ways and that we need flexible methodologies to understand and reflect on their influence on our social lives (Ziewitz, 2017). Examples are using combinations of qualitative and quantitative tactics, creative explorations (such as algorithmic or data walks), and other interdisciplinary approaches (e.g. Kitchen et al., 2014; Hyysalo et al., 2019; Ziewitz, 2017).

Seaver (2017) highlighted multi-sitedness and flexible methodologies by arguing that ethnographers should take a “scavenging approach” to study data-driven technologies. He said that we should see ourselves as eclectic “scavengers” moving from site to site, using mobile approaches to collect data drawn from disparate sources. A scavenger traces an object as it travels and is enacted across various sites. The scavenger “replicates the partiality of ordinary conditions for knowing – everyone is figuring out their world by piecing together heterogeneous clues – [and] expands on them by tracing cultural practices across multiple locations” (Seaver, 2017: 6-7).

I found the idea of scavenging eminently suitable for this research project. In the wild, scavengers feed themselves partly or wholly on decaying bodies. Therefore, they need to be alert, look around and pick up all sorts of traces to find something to eat. Much like foragers, they collect scraps, the leftovers and use them to assemble their next dinner, instead of hunting down other animals. To succeed, scavengers need to be flexible in what they eat and good at adapting to new environments compared to other organisms. Scavengers have an important role in ecosystems as they break down all sorts of organic material and process all sorts of nutrients for others coming after them in the food chain.

Similarly, I sought traces of data-driven technologies and tried to collect heterogeneous clues in diverse locations. I sifted information on data-driven technologies from several conferences on Big Data and data science media, pursued online courses on machine learning, picked through informal conversations with friends, family, and people in the healthcare field more familiar with programming and data-driven technologies than myself, and discussed and experimented with algorithmic walks with my students.

Some of the scraps I found needed more attention and became my research sites. I traced the envisioning of data-driven technologies back to writing in scientific journals from the healthcare domain (Chapter 2) and in the accounts of experts involved in Big Data initiatives stimulated by the European Union (Chapters 3 and 4). I tracked down the construction of data-driven technologies to a pioneering hospital-based initiative in the Netherlands (Chapter 5) and again in a machine learning training program for medical professionals that stimulated the normalization of the technologies (Chapter 6).

To make sense of all the traces found in diverse healthcare settings, I had to be flexible and adapt my research methods to the setting. I used a combination of qualitative methods, depending on what was suitable and possible in each research site. The data collection lasted from December 2016 to July 2019 and involved a total of 164 interviews combined with more than 250 hours of observations and extensive document analyses. When necessary, I also experimented with innovative methods of data analysis. For example, I developed a text mining tool to analyze a large amount of data in Chapter 4. As a scavenger, I did not collect and analyzed all these data alone; scavengers often work in packs or teams. Similarly, all the empirical chapters were written collaboratively with my co-authors, with whom I frequently discussed the data.

## Outline of the chapters

In the following chapters, I assemble healthcare practices that are reconfiguring responsible knowledge practices in response to data-driven technologies.

**Chapter 2** traces data-driven technologies back to the editorials of scholarly journals in the healthcare domain. Here, we study how Big Data is **perceived** by identifying which epistemic discourses are influential in envisioning Big Data. The chapter is based on a systematic literature review and gives insight into how Big Data use is validated, reinforced and its epistemic superiority is claimed. It highlights five discourses that frame data-driven technologies. Three discourses (the modernist, instrumentalist and pragmatist) disseminate a compelling rhetoric that presumes that Big Data are benign and lead to valid knowledge. The two other discourses (the scientist and critical-interpretative) question the objectivity and effectivity claims of Big Data, but are in the minority.

**Chapter 3** focuses on the ethical **framing** of data-driven technologies in envisioning data-driven healthcare. It is based on 137 interviews with diverse experts involved in Big Data initiatives in eight European countries as well as document analyses. The chapter identifies three forms of ethical framing: ethics as a balancing act, as a technical fix and as a collective thought process. Each way assigns roles and responsibilities to various actors in order to create responsible knowledge practices.

**Chapter 4** uses the metaphor of dreams to gain insight into how experts in Big Data initiatives envision data-driven technologies to improve knowledge practices in healthcare. Again, this chapter is based on the 137 interviews and document analyses mentioned above. It describes how experts **dream** that data-driven technologies can help overcome



general, scattered, slow and uncontrollable information in healthcare and gives insights into the experts' motivations, values and considerations.

**Chapter 5** is based on six months of ethnographic fieldwork in a Dutch hospital-based data-driven initiative in psychiatry. It observes how medical practitioners invited data scientists to construct prediction models of patient outcomes based on machine learning techniques. It analyzes the differences in epistemic culture and shows how data scientists and medical practitioners **negotiate** on epistemic virtues to create responsible knowledge practices.

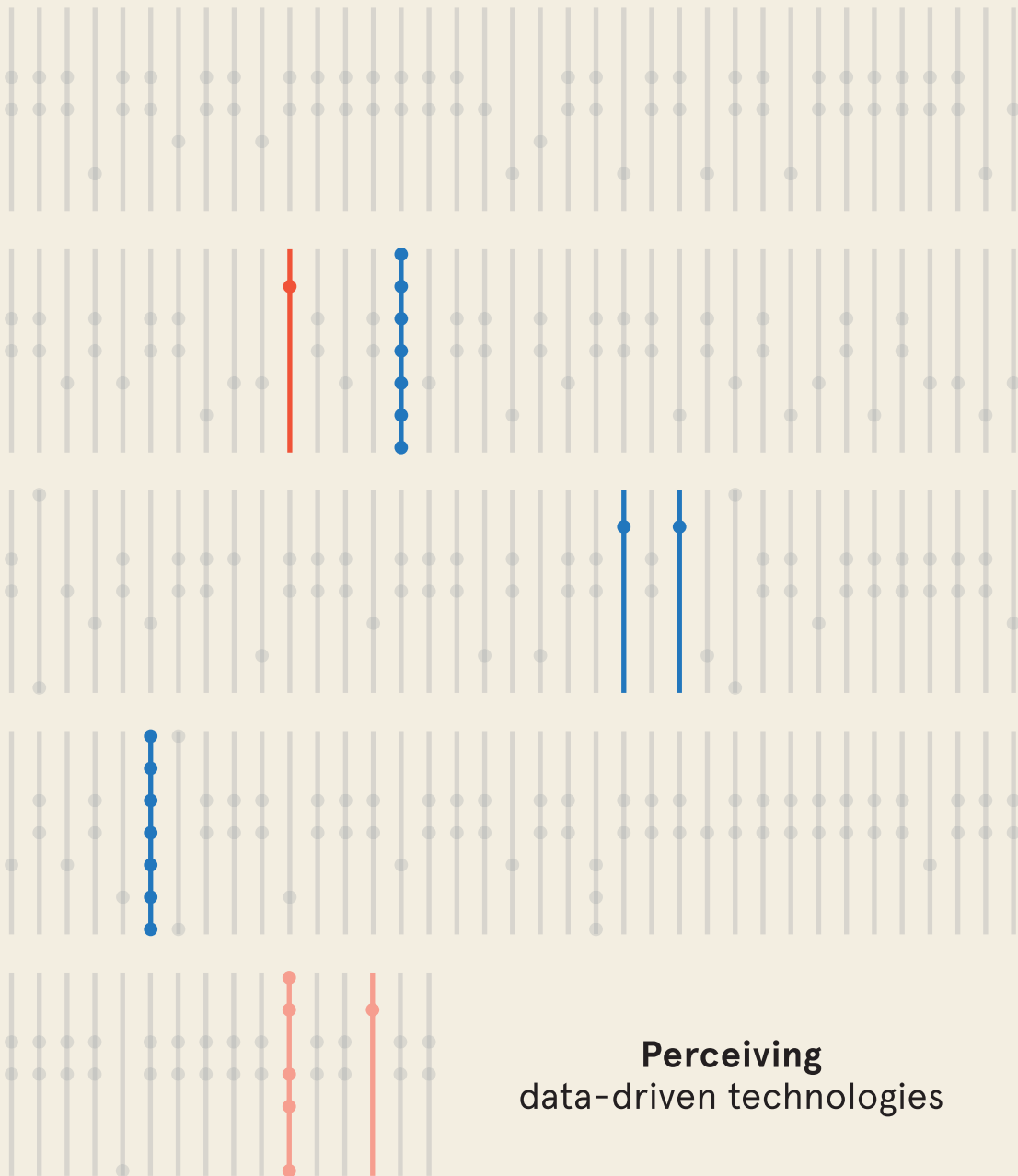
**Chapter 6** focuses on epistemic responsibility-in-the-making. The chapter is based on ethnographic fieldwork observing 14 Dutch mental healthcare professionals studying the basics of machine learning during a four-month course while pursuing a machine learning project in their own organizations. The chapter draws upon feminist literature on care to study how both the technology and the medical professionals **care** for responsible knowledge production.

**Chapter 7** presents my conclusions. I **reflect** on the various healthcare settings studied in the chapters and answer the research questions presented in this introduction. I outline the theoretical, practical, methodological implications and, finally, suggest a research agenda for future research on data-driven technologies.

Published as:

Stevens M, Wehrens R and de Bont A (2018) Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society* 5(2): 1–21.

# Chapter 2



**Perceiving**  
data-driven technologies



## Introduction

In recent years, the healthcare sector has welcomed an emerging field of practices captured under the umbrella term of 'Big Data'<sup>1</sup>. Big Data initiatives are welcomed because of their envisioned benefits for faster and more representative knowledge that is presumed to improve the process, management and predictability of care (Murdoch and Detsky, 2013). The healthcare sector traditionally favors high-quality evidence from randomized controlled trials (RCTs) and observational studies to guide treatment decisions and to organize the field (Timmermans and Berg, 2003). However, as the persistent discussions about evidence-based medicine show, the field has been struggling with the reductionist and generalized character of this evidence (Berwick, 2016; Greenhalgh et al., 2014). Patient guidelines are, for example, often based on time-consuming RCTs and done on selective populations, which makes it hard to extrapolate results to individual patients (Felder and Meerding, 2017). Big Data seem to offer an attractive alternative and are surrounded by claims of quick and comprehensive analysis of data and "with the aura of truth, objectivity and accuracy" (boyd and Crawford, 2012: 663).

Publications about Big Data frequently discuss topics related to knowledge generation, evidence and causation (e.g. Anderson, 2008; Mayer-Schönberger and Cukier, 2014). Provocatively, these publications celebrate the inevitable decline of traditional research as Big Data are supposed to handle large volumes of messy real-world data more efficiently and can uncover hidden correlations. In response to these claims, there has been a recurrent call for more studies into the epistemological implications of Big Data (boyd and Crawford, 2012; Crawford et al., 2014; Mittelstadt and Floridi, 2016), which scholars have started to address. As a result, a critical scholarly discourse that reflects on how Big Data shape our knowledge and understanding is forming in, primarily, the fields of STS and CDS (e.g. Kitchin, 2014; Leonelli, 2014; Rieder and Simon, 2016). While these fields have been instrumental in elaborating the neglected and problematic dimensions of Big Data, it remains an open question how and to what extent such insights become embedded in other fields, such as healthcare.

This chapter critically reviews the epistemological claims and envisioned implications that accompany Big Data in healthcare. The healthcare sector is, in general, characterized by a strongly institutionalized set of epistemological principles and generally accepted scientific methodologies (Timmermans and Berg, 2003). Big Data challenge these principles and methodologies with the consequence that the epistemological implications of Big Data practices could be particularly profound. What we value as evidence and knowledge has implications for the way medical decisions are taken and healthcare is organized. Opening

up the assumptions allows us to evaluate the role of Big Data in healthcare critically and open up opportunities for debate and fruitful intervention.

We base the chapter on a systematic and comprehensive review of scientific editorials as these, in particular, summarize and reflect upon developments in the field. We focus on discourses surrounding Big Data in the analysis and construct five ideal typical discourses based on a detailed analysis of the language conveyed in the editorials. The discourses show the diverse ways in which Big Data and the epistemological claims are conceptualized. We chose this focus as language is the medium through which people come to understand Big Data and it influences the way Big Data initiatives are performed and legitimated. Three questions guide our analysis: (1) *What Big Data discourses can be identified in scientific healthcare literature?* (2) *How do the discourses conceptualize the meaning of evidence?* (3) *What are the consequences of these conceptualizations for the way Big Data is understood in healthcare?*

## Big Data as material practice and semantic reality

Many authors have discussed the ambiguity surrounding the term Big Data. The term is often characterized by its volume, velocity and variety ('the 3Vs', Mayer-Schönberger and Cukier, 2014). However, many believe that these three characteristics do not sufficiently capture Big Data. The 3Vs are thus often extended with extra 'V's, such as value, viability, variability, visualization and veracity (DeVan, 2016; Kitchin and McArdle, 2016). Others use different qualifications to characterize Big Data, such as exhaustively, relationality, extensionality and scalability (boyd and Crawford, 2012; Kitchin and McArdle, 2016; Mayer-Schönberger and Cukier, 2014). Despite the many attempts, there is still no consensus about the term Big Data.

Inspired by the approach of Beer (2016) and Rudinow Saetnan and colleagues (2018), we conceptualize Big Data as a set of practices and ideas that exist in both (1) real material practice and in (2) a semantic reality. First, Big Data exist in specific actions, technologies and initiatives that are introduced to restructure healthcare. It is linked to the collection and aggregation of available data and correlation, pattern-recognition and predictive analyses. These data and analytics are subsequently used in real initiatives that aim to collect data, track, profile and predict behavior, preferences and characteristics (Mittelstadt and Floridi, 2016). Second, Big Data exist in a semantic reality as it is something that we talk and write about in order to anticipate the (possible) effects. In this semantic reality, we envision and give meaning to the present and future of Big Data. Of course, the way we describe Big

Data subsequently influences the way Big Data are performed and legitimated and vice versa.

In this chapter and our analysis, we focus on the semantic reality of Big Data and discourses and metaphors. This is not to argue that detailed empirical investigations into material practices are less important. However, if we want to explore the implications of Big Data, we also need a better understanding of how Big Data are discursively constructed. The crucial role of metaphors<sup>2</sup> in people's experience and sensemaking of the world has been long recognized (Lakoff and Johnson, 2011) as metaphors play a large role in framing debates in particular ways. Metaphors are not neutral as they embody assumptions, imagined implications and impose opportunities and limitations (Puschmann and Burgess, 2014; Zinken et al., 2008). This makes metaphors especially valuable as we want to open up the epistemological claims and assumptions that accompany Big Data in healthcare.

## Methods

We conducted a comprehensive and systematic search of scientific literature to show the different ways in which Big Data and its epistemological claims are being articulated in the healthcare sector. We chose this approach because we did not want to miss major views and also gain insight into the relative spread of the articulations. Although our search of the literature fits the methodological approach of a systematic literature review, we subsequently departed from this approach in the interpretation and analysis of the results. While a 'traditional' review counts and synthesizes the results and provides an exhaustive summary of current evidence, we chose to follow a discourse analytic approach for the analysis because we wanted to move beyond a summary of results to provide an interpretation of the material (Dixon-Wood et al., 2006). The main advantage of this approach is that it combines the strengths of a systematic, thorough literature search with the explanatory power of interpretive analyses that provides new insights into a phenomenon.

### Identifying relevant studies

A search term was composed with the help of a librarian to select the relevant studies. The search term covered terms related to (1) 'healthcare' and (2) 'Big Data' and related techniques, such as data mining. We wanted to be as inclusive as possible. The librarian and the first author looked for mentioning of the term Big Data in relevant studies and included those. Also, they started with a small list of techniques related to Big Data and iteratively added additional techniques to the search term if they were frequently mentioned in the found studies and resulted in relevant studies. The minimum requirement for inclusion was the mentioning of unusually large data sets or combinations of diverse

types of data sets. We choose not to include the search term 'artificial intelligence' as this resulted in thousands of studies more for inclusion. In addition, we decided not to include 'knowledge', 'evidence' and related terms in the search profile because we assumed that even studies that do not mention these terms can still make epistemological claims. The exact search terms can be found online <sup>3</sup>. Eventually, we conducted the extensive search in Embase, Medline Ovid, Web of Science, Scopus, LISTA EBSCOhost and Google Scholar in January 2017.

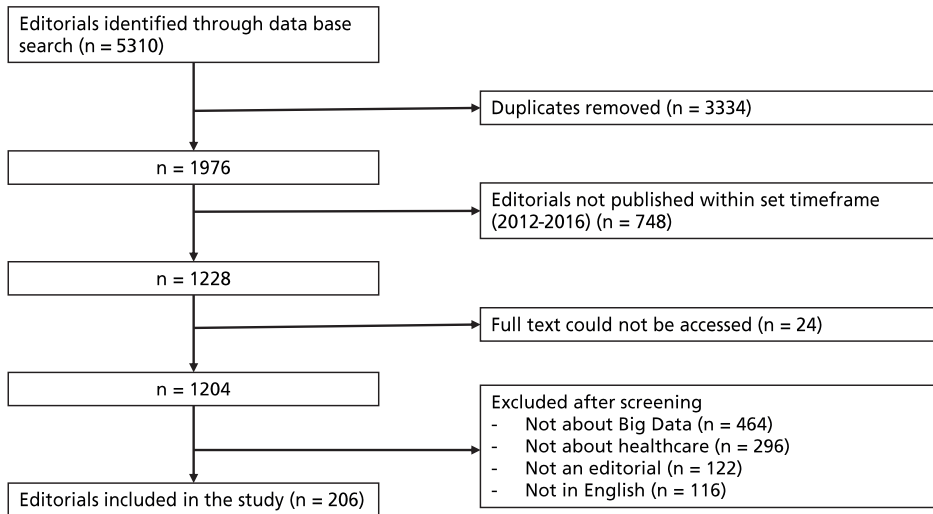
We chose to limit our search to editorials from scientific journals in the healthcare domain because of their distinct characteristics. Editorials are expressions, reflections, or commentaries on developments. They are a medium for editors, researchers and clinicians to communicate with peers and informed publics, as well as a forum for the explicit expression of beliefs and opinions (Loke and Derry, 2003; Miller et al., 2006). They can contain substantial scientific content, compelling messages, calls for action and discuss little known scientific facts with far-reaching consequences (Rousseau 2009). They are usually written by the journals' editors or leading authors of the field. Editorials are often accessed and appear in well-regarded academic journals (Loke and Derry, 2003; Youtie et al., 2016). We selected editorials instead of viewpoints and opinion articles because we assume that editorials have a more critical role in defining the standpoint of the journal as compared to presenting the opinions of individuals. Lastly, editorials set the agenda for specific research fields and are a basis for future action. Hence, we believe that editorials capture Big Data discourses in the scientific community and have an important function in disseminating assumptions about Big Data in the healthcare domain.

Given the size of the original body of selected documents, further selection criteria were needed to obtain a manageable data set for detailed analysis. Hence, we chose to define a timeframe (2012–2016) for the review. As other studies have, we noticed an exponential increase in the number of publications about Big Data in general in 2012 (Youtie et al., 2016). Therefore, we choose 2012 as the starting point. Also, we included only English language editorials for practical reasons. If we could not find the editorial text online, we contacted the first author to gain access. In 24 instances, this did not work, and these documents were excluded because we could not access the full text.

The final selection of documents contained 1204 original documents. The first author of this chapter read the title and abstract or the first and last paragraphs (if an abstract was unavailable) and excluded the irrelevant texts. Documents were excluded in close cooperation with the second and third authors because they either did not qualify as editorials or were outside the scope of this review (i.e. documents that were not about Big Data or were unrelated to health or healthcare). After screening, 206 editorials were eventually



included for detailed review (see also Figure 1). An overview of the included editorials can be found online <sup>3</sup>.



**Figure 1.** Selection of the editorials

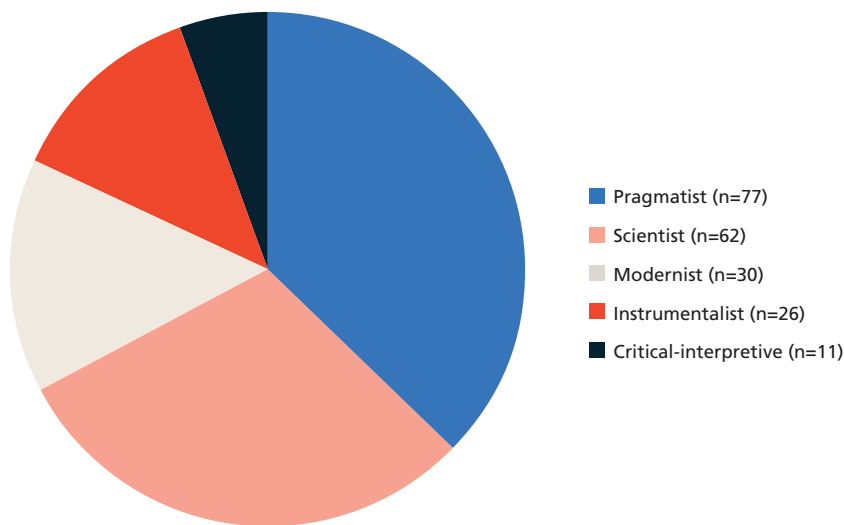
## Data analysis

The analysis was conducted in three phases. First, the first author randomly selected 20 editorials and flagged sections of interest. The authors of the chapter discussed trends in the editorials and composed a list of questions that would be relevant to answer for each editorial. Subsequently, the first and second authors analyzed another 20 editorials and the list of questions was finalized. The list contained questions about (1) conceptualization of Big Data (e.g. how is Big Data described?), (2) the epistemological position (e.g. what is described as a good way of obtaining evidence/knowledge?), (3) the envisioned consequences (e.g. how are outcomes of Big Data used?) and (4) noticeable discursive elements, such as metaphors and surprising examples or comparisons. In the second phase, all remaining editorials were analyzed with the finalized analytical scheme by the first author, second author and a junior researcher. The questions were answered for all the editorials and organized in a spreadsheet. Ten percent of the editorials were also analyzed by another member of the research team to ensure analytical consistency. Third, to organize and interpret the spreadsheet and to construct the ideal typical discourses, the authors of this chapter jointly tested, critically interrogated and experimented with the analytical themes and organization of results until consensus was reached about the structure and characteristics of the several discourses. This process eventually resulted in the construction of the five discourses.

## Results

### Description of the data set and overview of findings

Based on our analysis, we were able to construct five ideal typical discourses: modernist, instrumentalist, pragmatist, scientist and critical-interpretive. We drew inspiration for the names of the discourses from the relations we saw between implicit assumptions about evidence and knowledge and diverse philosophical and epistemological positions. The discourses were distributed over the editorials in the following way: modernist (n = 30), instrumentalist (n = 26), pragmatist (n = 77), scientist (n = 62) and critical-interpretive (n = 11; see Graph 1). These discourses should be viewed as ideal types, meaning that some editorials consist of combinations of various discourses. Co-occurrence especially consisted between the instrumentalist and pragmatist discourses (n = 16) and between the modernist and pragmatist discourses (n = 12). The modernist and critical-interpretive discourses and the instrumentalist and critical-interpretive discourses did never co-occur in one editorial.



**Graph 1.** Presence of the ideal typical discourses in the editorials

We summarized the discourses and their main characteristics in Table 1. We will describe the five ideal typical discourses in more detail below. In our description of the discourses, we will highlight one metaphor that is particularly apt to illustrate the epistemological positions of each specific discourse.

Ideal typical discourse	Modernist	Instrumentalist	Pragmatist	Scientist	Critical-interpretive
<b>Big Data is described as</b>	Large amounts of data that can be analyzed	Analytic techniques	A useful (managerial) instrument for decision-making	A trend that deals with data collection, analysis and outcomes more flexibly	A trend that oversimplifies reality
<b>Evaluation of Big Data</b>	Positive	Positive	Positive	Critical	Critical
<b>Recommendations for further development</b>	Start to use Big Data in healthcare	Enhance and develop the Big Data techniques	Implement Big Data in healthcare	Be (extremely) careful with the use of Big Data	Discuss the negative consequences of Big Data
<b>Non-use of Big Data is explained in terms of</b>	Not discussed	Techniques do not work sufficiently	Implementation problems	Lack of performance (as traditional studies perform better)	Negative consequences for individuals and society
<b>Inference from data</b>	Direct, data equals knowledge	Direct, data equals knowledge (that we can see through the techniques)	Direct, data equals knowledge (if useful in practice)	Indirect, data interpretation involves scientific methodology (hypothesis testing)	Indirect, data interpretation involves critical thinking
<b>Epistemological claim</b>	Big Data offers reliable information	Big Data offers increasingly more reliable information as the techniques improve	Big Data can offer reliable information in some situations	Big Data can be useful if strict criteria are met.	Big Data will always generate limited evidence
<b>Presumed reliability of Big Data</b>	High	High	High	Medium - Low	Low
<b>Summarizing metaphor</b>	Capturing data	Illuminating data	Harnessing data	Selecting data	Constructing data
<b>Presumed consequences of Big Data</b>	Revolutionary amount of new knowledge	New predictions and increased understanding to solve persistent problems	Improved problem-solving and decision-making in healthcare	Inconclusive and misguided information, if Big Data are not properly used	Inconclusive and misguided information and unfair outcomes

Table 1. Overview of the discourses

## The modernist discourse: Capturing data

### *The conceptualization of Big Data*

Big Data are often not defined in this ideal type, but the editorials link it to large amounts of data. Big Data are described as a positive development and the editorials stress the beneficial effects of Big Data. They state, for example, that it will lead to proactive, predictive, preventive, participatory and patient-centered health (Shah and Tenenbaum, 2012; Weinstein, 2016). However, the precise meaning of these statements often remains unclear and ambiguous, as they are not discussed further.

The editorials unanimously and unambiguously recommend the use of Big Data in healthcare. This is emphasized by three rhetorical techniques. First, these editorials' tone is optimistic, signified by such words as 'explosion', 'revolutionizing', and 'world-changing possibilities'. Big Data are presented as innovative and as a rupture with the past that will radically transform healthcare (Restifo, 2013; Weinstein, 2016). Secondly, a sense of urgency is created in the editorials as they often draw a contrast between the medical domain and other sectors that supposedly already take advantage of Big Data. The medical domain is presented as slow, conservative and old-fashioned, while other domains are already taking Big Data analytics for granted. This discursively constructs the field of medicine and its current approaches as unsustainable and outdated (MacRae, 2012; Risoud et al., 2016). Third, there is almost no attention to the negative sides of Big Data, such as potential issues with privacy, consequences of shifting power-relations, or practical questions concerning implementation. Illustrative of this position is the almost complete lack of non-use of Big Data as a theme in this discourse.

### *Epistemological assumptions*

Capturing data is the metaphor (Figure 2) that most clearly illustrates the epistemological assumptions in the modernist discourse. First, because the modernist discourse assumes data to exist in the world and to have inherent value (like a butterfly or other natural resources). The assumptions are that the data can be captured and that this results in new insights, evidence and practices. Second, the metaphor aptly illustrates the epistemological assumptions in this discourse because capturing is a relatively simple act that also leaves the data itself unaffected, which shows the ease in which Big Data are portrayed in these editorials to be able to arrive at knowledge. This process is viewed in such simplistic terms that data seem to equal knowledge. This creates the idea that only 'capturing data' already leads to new knowledge.



Figure 2. Capturing data metaphor

### *Consequences*

The modernist discourse strives for a radical change as the traditional ways of knowledge production in the medical domain are rejected. Editorials in the modernist discourse aim to overthrow the status quo in order to transform knowledge production in healthcare radically. Big Data are seen as a legitimate source of knowledge in these editorials because Big Data are argued to lead to more timely and reliable knowledge that is viewed as immediately useful in practice. However, the discourse seems to be naïve in the sense that it only addresses grand visions and is not concerned with, for example, the practical development and application of Big Data, nor with the societal effects.

## **The instrumentalist discourse: Illuminating data**

### *The conceptualization of Big Data*

In this ideal type, Big Data are understood in terms of a range of analytical techniques, such as pattern recognition, data mining and machine learning (Amato et al., 2013). The editorials have a positive tone and describe ways in which these Big Data techniques can aid

healthcare, for example, by predicting disease outcomes and increasing the understanding of the causes of diseases (Belgrave et al., 2014; Van De Ville and Lee, 2012). The editorials typically discuss how analytic techniques should be used and how they can be improved. The editorials contain advice on how one should deal with the missing data, correlated features and replication and separation of training and validation sets.

The editorials recommend that Big Data techniques should be developed and enhanced to gain better results. Editorials in this discourse place a high value on experimentation. For example, innovative studies in which Big Data techniques are used for brain decoding and the development of clinical decision support systems are presented (Najarian et al., 2013; Van De Ville and Lee, 2012). Using Big Data techniques for these purposes is by no means standard practice, but by trying out and experimenting with data analytic processes, the techniques are improved. Illustratively, terms like improving, experimenting, exploring, developing and learning frequently occur in the instrumentalist editorials.

### *Epistemological assumptions*

The illuminating data metaphor (Figure 3) best represents the epistemological assumptions in the instrumentalist discourse and is exemplified by phrases such as ‘casting light’ and ‘highlighting’ in the editorials. Similar to the modernist discourse, in the instrumentalist discourse data seem to exist in the world and are viewed as having an intrinsic value. However, the process of knowledge discovery through Big Data is depicted in less simplistic terms than in the modernist discourse, as the editorials emphasize that information can only be extracted from highlighting the data with specific analytic techniques so that patterns in the data can be seen (Amato et al., 2013; Rosenstein et al., 2014). This is an indirect critique of the more traditional methods for knowledge generation, which are implicitly depicted as outdated and inefficient. The editorials thus suggest that by constructing and positioning the ‘light sources’ (e.g. the analytic techniques), we are increasingly able to ‘see’ the data and emerging trends within them. This means that knowledge improves together with the set of analytical techniques.

### *Consequences*

The instrumentalist discourse promotes the use of Big Data techniques in healthcare as they become a reliable source for decision-making. Less radically than the modernist discourse, editorials in this discourse still argue for a change of the ways knowledge is obtained in healthcare, as Big Data are expected to solve persistent problems in healthcare. The discourse seems to envision Big Data as a tool to solve problems and the tool is valid to the extent that it helps make accurate predictions and increases our understanding. However, similar to the modernist discourse, the instrumentalist discourse also neglects the broader implications and potential societal effects of the use of Big Data techniques.

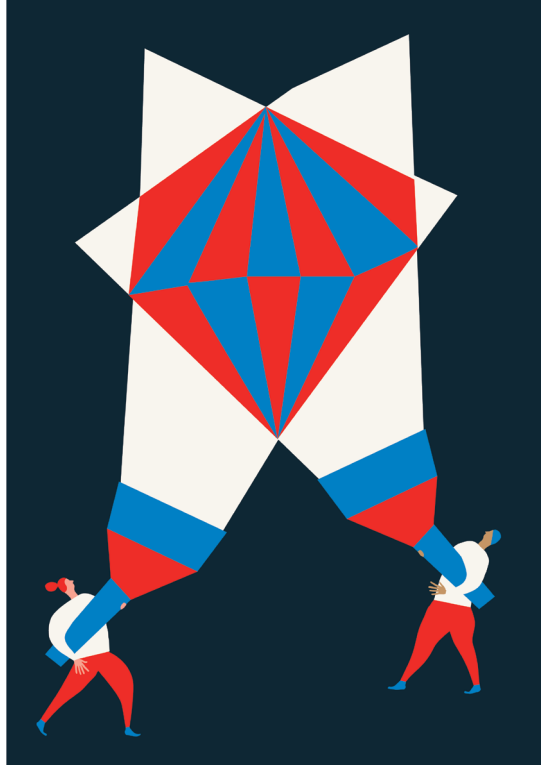


Figure 3. Illuminating data metaphor

## The pragmatist discourse: Harnessing data

### *The conceptualization of Big Data*

In this ideal type, Big Data are conceptualized as a useful (managerial) instrument for problem-solving and decision-making in healthcare (Garrison, 2013; Klonoff, 2013; Potters et al., 2016). Big Data are discursively constructed in the editorials as a phenomenon that is already here and is likely to stay (Basak et al., 2015; Ghani et al., 2014; Hay et al., 2013). Big Data are described as a positive development. However, in this discourse, people are presumed to have a significant influence on how Big Data take shape, as opposed to the more technological determinist pattern of thinking that characterizes the modernist discourse.

The editorials in this discourse primarily focus on how Big Data should be implemented and describe the steps for successful implementation. They discuss, for example, the training, recruitment and the introduction of data scientists or knowledge engineers, cultural factors that need to change in healthcare, new rules and regulations that have to be made, the

adoption of new platforms and information systems, and how access should be gained to the data and analytics (Cases et al., 2013; Kottyan et al., 2015; Narula, 2013; Potters et al., 2016). The editorials do mention concerns and other challenges that need to be overcome or solved, as the following quote from McNutt et al. (2016: 914) illustrates:

*“We envision future systems that incorporate [Big Data] decision support models into the clinical systems in ways that enable clinicians to improve both the quality and the safety of care they give and the efficiency with which they give it. To reach this vision, there remain technological needs and human challenges to overcome.”*

### *Epistemological assumptions*

The metaphor of ‘harnessing data’ (Figure 4) best illustrates the ideas and assumptions about Big Data in the pragmatist discourse. Similar to the previous discourses, data continue to be described as something ‘out there’, simply existing in the world. The data are viewed as valuable as they can be translated into information and knowledge. Different is that this discourse sees traditional scientific and Big Data methods as complementary approaches that can both generate ‘evidence’ and have practical relevance (Basak et al., 2015; Klonoff, 2013). A more pragmatic attitude towards evidence seems dominant as evidence is not strictly related to scientific processes. There are no fundamental objections against using Big Data outcomes. Big Data are viewed as beneficial whenever it helps to gain knowledge about situations that traditional scientific methods cannot study and decision-makers pragmatically make choices based on the available evidence. Discussions about the status of the outcomes of traditional scientific studies and Big Data analyses disappear to the background in this discourse, as the actionable character is emphasized.

### *Consequences*

Similar to the instrumentalist discourse, the pragmatist discourse envisions a change in the way decisions are taken as Big Data offer more knowledge than currently is available and can generate useful new insights for healthcare practice. Big Data are seen as a valuable source for decision-making next to traditional knowledge-producing approaches. This discourse deals – more than the previous discourses – with some of the practical issues surrounding Big Data implementation (such as the recruitment of data scientists). However, the epistemological and normative changes that Big Data bring are not addressed.

## **The scientist discourse: Selecting data**

### *The conceptualization of Big Data*

In this ideal type, Big Data are described as a new trend that deals with data collection, analysis and outcomes in a less rigorous way than scientific methodologies do. The edi-





**Figure 4.** Harnessing data metaphor

torials mention that Big Data can be useful in some situations because of its potential to identify valuable research directions, for hypothesis-generation and exploration of massive data sets (Khoury and Ioannidis, 2014; Krakoff and Phillips, 2016). It can thus only be used as exploratory, hinting at possible directions for traditional research designs. The tone of the editorials is critical, especially compared with the modernist discourse, and Big Data are seen as a potentially dangerous development.

The editorials argue for caution with regards to Big Data and claim that traditional scientific methods will remain essential despite the arrival of Big Data methodologies. The editorials try to distinguish 'proper' from erroneous science. For example, they do this by comparing Big Data outcomes and findings from RCTs (Freeman and Saxon, 2015). Some editorials mention the limitations of traditional studies. For example, they state that RCTs are costly or not always possible because of ethical considerations (Freeman and Saxon, 2015; Leem, 2016). However, the consensus seems to be that despite the potential of Big Data as a starting point for research, it always needs to be followed by more substantive research.

Or as Khoury and Ioannidis (2014: 1054) state in their editorial: “We should embrace (and not run away from) principles of evidence-based medicine.”

### *Epistemological assumptions*

The epistemological assumptions about Big Data within this discourse can be summarized by the metaphor of ‘selecting data’ (Figure 5). The notion that Big Data can lead to reliable and valid knowledge is questioned and sometimes outright denied in the editorials. Two arguments are frequently made. First, the editorials stress that data are essential to arrive at knowledge. However, data are not viewed as pre-existing in the world. As such, they cannot simply be captured, illuminated or harnessed, but need to be selected and processed via specific methods. This position is reinforced by statements like ‘garbage in, garbage out’ (denoting the idea that the lack of selecting ‘high-quality’ data from the masses of available, often poor quality data leads to useless analyses), or by presenting the data of Big Data as erroneous or as a ‘dumping site’ (Brown, 2016; Patrick, 2016). Through discursively oppositioning high-quality data with ‘garbage’, the editorials point to the need to have the proper or right procedures for data gathering and analysis in place. Such procedures are meticulous and less easily abandoned than presumed in, for example, the modernist discourse. Second, the editorials problematize the assumption that more data equal better knowledge. This idea is widespread in the modernist, instrumentalist and – to some extent – pragmatist discourses. According to editorials in the scientist discourse, this assumption is wrong. As Onukwugha (2016: 92) explains:

*“We cannot assume that more data necessarily means more information. Indeed, as the volume of data increases, it will be important to pay continued (or more) attention to established concerns regarding measurement, bias, and fallacies relevant to empirical analysis and interpretation.”*

Despite the criticism, the epistemological position is similar to the modernist and instrumentalist discourses as the positivistic notion that truth can be found in data is also present. However, in the modernist and – to some extent – instrumentalist discourse there seem to be an acceptance of a rather naïve empiricism that, according to the scientist discourse is too simplistic. The scientist discourse argues that, for example, Big Data can be informative but never capture a whole domain and that there remains a need for hypotheses and theory. So, evidence is assumed to be developed only by correctly applying the scientific method. Just experimenting with Big Data can lead to wrong conclusions (Gomella, 2016).

### *Consequences*

The scientist discourse argues against a radical change in healthcare as, according to this discourse, Big Data are not a reliable source of knowledge. The only proper knowledge



Figure 5. Selecting data metaphor

seems to be scientific knowledge and such knowledge can only come from the use of strict scientific methods. The consequences of Big Data would be false evidence and knowledge with possibly large, detrimental effects. This discourse discusses in-depth the epistemological concerns and how Big Data related to traditional structures for knowledge generation.

## The critical-interpretive discourse: Constructing data

### *The conceptualization of Big Data*

In this ideal type, Big Data and data are presented as an oversimplified presentation of reality. The critical-interpretive discourse incorporates diverse forms of criticisms. Generally, the editorials share a concerned tone and their criticisms are both epistemological and societal.

The editorials advocate discussion on the position of Big Data in our society as a whole. Two lines of critique can be distinguished in this discourse. First, the simplicity of data is frequently addressed. Big Data are dismissed because it is a reductionist and oversimplified presentation of reality, unable to adequately capture and account for the richness

and diversity of human experience. Editorials make this point by describing data that are missing in Big Data sets and by stressing the importance of personal experience, objectives and preferences (Pope et al., 2014; von Gunten et al., 2016; Zurlinden, 2016). Second, the editorials stress the normative aspects of Big Data and point out that these aspects are often overlooked or neglected. For example, the editorials focus on the danger of Big Data that is not being interpreted by physicians and warn that Big Data can be a first step for 'dangerous' automatic decision models. As Von Gunten et al. (2016: 1240) state: "It [Big Data outcomes] must be interpreted by a seasoned clinician with critical thinking skills."

### *Epistemological assumptions*

The epistemological assumptions that characterize editorials in this discourse can be best understood via the metaphor of 'constructing data' (Figure 6). In terms of epistemological assumptions, the critical-interpretive discourse is most distinctive from the other discourses as it reasons from a different set of epistemological assumptions (building on constructivist traditions in philosophy of science as opposed to positivist approaches). Consequentially, data are no longer presented as something given that can be captured or illuminated



**Figure 6.** Constructing data metaphor

but understood as the result of the social and political processes that created them. As Pope et al. (2014: 68) state: “We must remember that all data – big or small – are socially constructed.” This perspective means a recognition that data always emphasize certain aspects of the world while leaving out other elements. Importantly, the constructed data present an image, but editorials in this discourse warn that this image can never be complete. This discourse can especially be contrasted with the modernist discourse, in which the ideal of ‘complete knowledge’ is maintained. Big Data, therefore, according to the critical-interpretive discourse, will always generate limited knowledge and data have to be handled with care.

### *Consequences*

The critical-interpretive discourse warns for the limitations of Big Data. According to this discourse, while Big Data create new possibilities for generating knowledge, the use of these possibilities is not seen as a positive change. The starting point is that it is better not to use Big Data (or at most only with great restraint). The consequences of Big Data would be that limited data are extrapolated and would lead to erroneous outcomes that could cause harm to people and healthcare systems. In addition, if people are not able to recognize the fact that data are constructed, for example, by the use of automated decision models, essential aspects of care would be lost.

## **Discussion**

Reviewing literature is a first step in gaining a better understanding of the epistemological implications of Big Data in healthcare. Based on a systematic literature search and consecutive interpretive analysis, we constructed five ideal typical discourses of Big Data in healthcare. These five discourses all highlight particular aspects of Big Data, neglecting others, and thereby frame Big Data and its (epistemological) implications in specific ways. This study is vital because discourses and metaphors pre-structure the way that the material practices of Big Data take shape. As such, they are highly consequential in shaping current and future debates on Big Data. In this discussion, we will take the next step by drawing attention to the political dynamics of the discourses. We build on insights from STS and CDS to point to issues that have been ignored or neglected in the current construction of the Big Data debate in healthcare editorials. We end with suggestions for future research.

We noticed that the discourses that frame Big Data in positive terms (modernistic, instrumentalist and pragmatist) were more present in our empirical material ( $n = 133$ , 64.6%). These discourses seem to reinforce each other in the idea that Big Data result in valid knowledge and that massive data sets and predictive analytics reflect the truth. These

grand promises could explain the strong positive rhetoric that surrounds the term Big Data and that drives the implementation of Big Data initiatives in healthcare. The corresponding metaphors of capturing, illuminating and harnessing data all embody closely related epistemological expectations. Data are presented as benign, objective, an asset for organizations, and not something that should be questioned. The three discourses all view the advancement of Big Data into healthcare as inevitable (Mayer-Schönberger and Cukier, 2014; Murdoch and Detsky, 2013), with the instrumentalist discourse more concerned about the development of the analytic techniques and the pragmatist discourse more concerned about the implementation of Big Data.

The discourses that frame Big Data in more critical terms (scientist and critical-interpretive) were less present in the editorials ( $n = 73$ , 35.4%). They both challenge the objectivity, effectivity and serviceability claims that are dominant in the positive discourse, do not view Big Data as inevitable and pose alternative possibilities. This is important for healthcare, as they make sure we reflect on Big Data knowledge. However, both discourses do this from different implicit philosophical positions (positivist and constructivist). Their metaphors of selecting and constructing data illustrate another political message that frame Big Data as limited and claims that positive Big Data discourses obscure the often serious implications for expertise and evidence.

Especially editorials in the critical-interpretive discourse were limited ( $n = 11$ , 5.3%). This is an interesting observation in the light of the increased attention for the problematic assumptions and epistemological difficulties of Big Data in fields such as STS and CDS, often offering fundamental criticisms about the claims and expectations surrounding Big Data. For example, that although data may appear objective, they are still constructed through subject–technology interactions (boyd and Crawford, 2012; Dalton and Thatcher, 2014; Kitchin and Lauriault, 2014). An important conclusion that can be drawn from our analysis is that such work has not broadly infiltrated the domain of medical editorials.

We argue that the healthcare sector would benefit from a more prominent critical-interpretive discourse, as three important issues would be neglected (as they are not addressed by the other discourses): (1) the normative assessment of Big Data, for example, the role that automatic decision models should play in the doctors' office and issues related to data access and consent (Mittelstadt and Floridi, 2016). (2) Reflection on the situatedness of data. Data do not speak for themselves and we must remember that they are always an oversimplification of reality. Reflection on what particular aspects of a phenomenon are emphasized in the data and what aspects are occluded is therefore crucial (boyd and Crawford, 2012; Mittelstadt and Floridi, 2016). (3) The social and political processes that create Big Data. While Big Data and data may seem objective to many, they still are subjective

and contain biases and other limitations which should be opened up (boyd and Crawford, 2012). We believe that the pragmatist discourse deals with the first issues too pragmatically and the scientist discourse with the last issues too statically and without enough attention for the social dynamics. Subsequently, the healthcare sector would benefit from more critical reflection and intervention.

Based on this review, we stress that the epistemological discussion in healthcare needs to be developed further and that we have to find ways to better integrate aspects of the critical-interpretive discourse in the healthcare domain. Based on this chapter, we suggest the following directions for further research:

1. Further study into the five ideal typical discourses could provide important insights into the ways (and extent in which) similar discourses and dynamics are also noticeable in other disciplines. Quantitative approaches could investigate correlations between the background of editors/authors and the discourses they endorse.
2. As discourses are not only part of editorials but also of broader cultural discussions, future research could study the various ways in which the semantic realities of Big Data intersect with material practices and vice versa. Especially warranted are comparative studies that open up the ways Big Data are depicted in different cultural domains and the sociotechnical imaginaries (Jasanoff and Kim, 2015) in which these depictions are embedded.
3. Empirical reflections on the material practices of Big Data are warranted as well. Discourses and socio-technical imaginaries are still part of theoretical discussions, while at the same time, many Big Data initiatives are started in healthcare. Studying such initiatives ethnographically is likely to provide highly valuable insights into the dynamic encounters between data and healthcare.

To conclude, the fields of STS and CDS have been instrumental in opening up discussions about the epistemological and ethical implications of an emerging field of practices, captured under the umbrella term 'Big Data'. On the basis of this study, we have to conclude that these reflections have not been embedded in the healthcare sector in any substantial way. Based on a systematic analysis of scientific editorials, we constructed five ideal typical discourses to gain a better understanding of how Big Data are discursively constructed. We observed that editorials in the critical-interpretive discourse were limited (only 5.3%). We conclude that the healthcare sector would benefit from a more prominent critical-interpretive discourse since important reflections on the normativity and situatedness of Big Data, as well as the social and political processes that create Big Data, are not addressed by the other discourses.

## Footnotes

<sup>1</sup> We see Big Data as a set of practices and technologies that is discursively framed under the umbrella term 'Big Data'. We do not see Big Data as a coherent unity and therefore set Big Data in plural form.

<sup>2</sup> Two recent studies explored metaphors used to describe Big Data in popular mass media and business press. The first study by Puschman and Burgess (2014) recognizes two Big Data metaphors in mass media. Both dominant metaphors stress the idea that data accurately reflect nature, society and culture, and that the presented units (e.g., data) are comparable and the results are reproduced. The other study (Maiers, 2017) examined business press and noticed the frequent use of oriental metaphors. Maiers recognized a vertical direction in the metaphors (e.g. deep analytics, data mining, and drilling down) that suggest the assumption that by going deeper, more details, accuracy and precision can be found. We were surprised by the strength of the positivistic ideas related to these metaphors of Big Data because these are not only part of popular mass media and the business press, but are also actively embraced by many medical researchers and are recognizable in the editorials of renowned scientific journals.

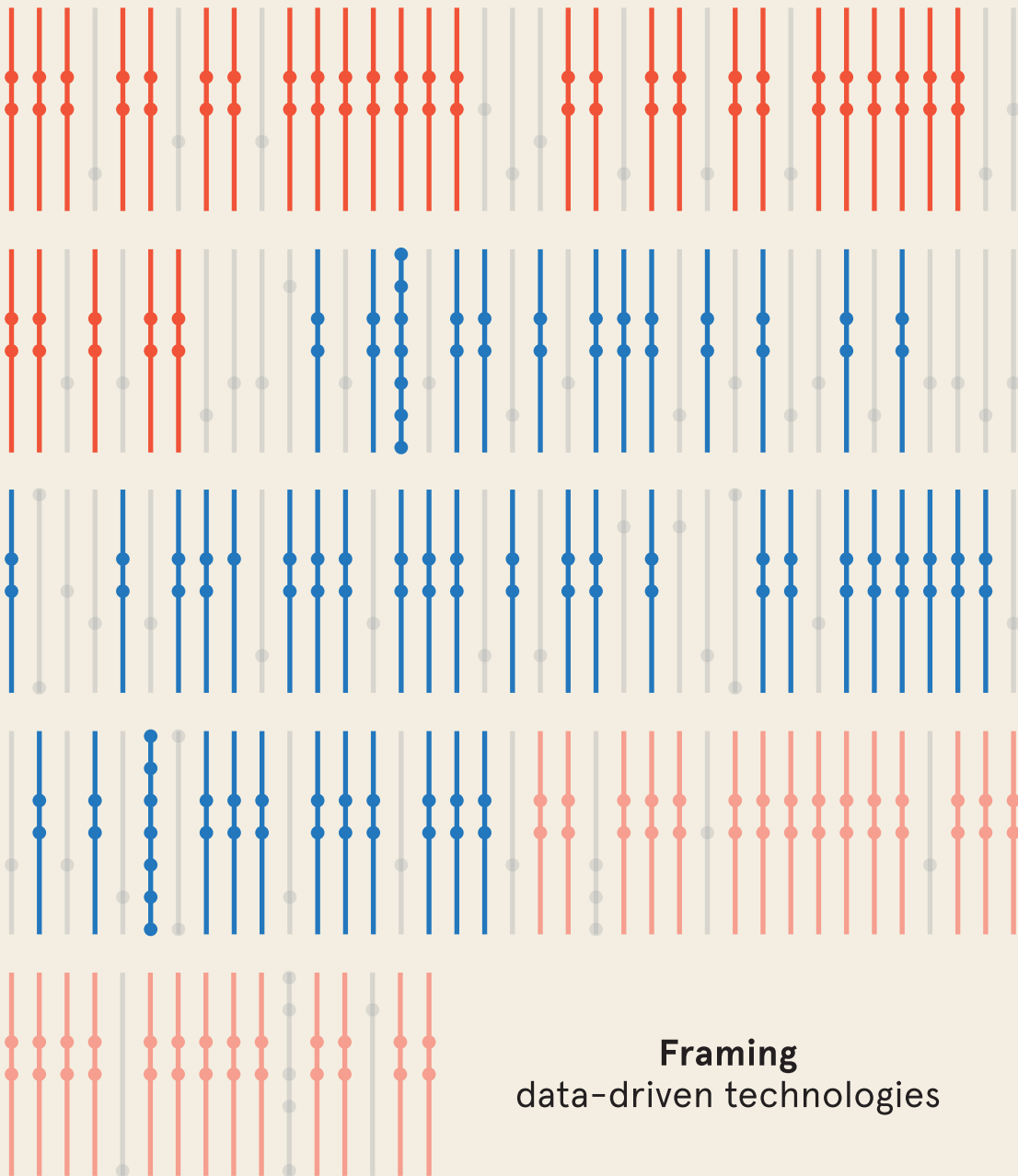
<sup>3</sup> These overviews can be found as attachments in the online version of the chapter. See: <https://journals.sagepub.com/doi/full/10.1177/2053951718816727>.





Under review at Science Technology, & Human Values as:  
Wehrens R, Stevens M, Kostenzer J, Weggelaar-Jansen AM and de Bont A. –  
“We can’t keep hiding behind the ethics”: The role of ethical framing in the  
promissory future of data-driven healthcare technologies.

# Chapter 3





## Introduction

The allure of a 'data-driven' future healthcare system continues to seduce many. Driven by technological developments (Mayer-Schönberger and Cukier, 2013) and increasing opportunities for algorithmic analysis of large and varied datasets (Raghupathi and Raghupathi, 2014), the latest years have seen a continuous stream of hypes and 'buzzwords' about the potential of data-driven healthcare technologies, ranging from Big Data, predictive medicine and data science to more recent expressions of hope and potential related to artificial intelligence and subfields like machine learning. The underlying promise is remarkably similar: healthcare is expected to improve significantly because more data are collected from varied sources and better analytical techniques are available to meaningfully process these data. Consequently, healthcare delivery is argued to become better, in terms of 'prevention' or 'personalization' (Raghupathi and Raghupathi, 2014; Kruse et al., 2016; Stevens et al., forthcoming). While experts in data science and artificial intelligence may shun from the conceptual indeterminacy of the above mentioned buzzwords in popular discourses, recent work in STS has analyzed how such buzzwords generate matters of concern, mobilize people by setting attractive goals, and steer the agenda through their 'promissory rhetoric' (Bensaude Vincent, 2014; Penkler et al., 2019). Indeed, such promissory rhetoric is highly recognizable in the perceived future of data-driven healthcare technologies (Hoeyer, 2019; Stevens et al., 2018).

An increasing strand of work in the fields of STS, CDS and CAS has started to interrogate the saliency of this promissory rhetoric by raising more fundamental ethical questions concerning epistemology, bias, surveillance and security, and opacity of data infrastructures (boyd and Crawford, 2012; Busch, 2017; Bauman and Lyon, 2013; Kitchin and Lauriault, 2014; Ebeling, 2016). Through this work, authors highlight the ethical dilemmas of data-driven technologies, providing a rich picture of ethical tensions, paradoxes and concerns (Mittelstadt and Floridi, 2016; Salerno et al., 2017; Winter and Davidson, 2019; Mittelstadt, 2019). An aspect that is less visible in this body of work, however, is how ethical arguments are used as part of discursive work by various practitioners engaged in data-driven initiatives in healthcare. Rather than a set of abstract principles, ethics can also be viewed as an emic resource that actors draw upon to distribute roles and responsibilities, legitimize particular courses of action, and negotiate the tensions and dilemmas involved in translating ambiguous legal principles into responsible courses of action (cf. Cool, 2019; Hoeyer et al., 2017).

This chapter argues for more explicit attention to the role of ethical framing in shaping the promissory future of data-driven healthcare technologies. We propose that studying how actors - engaged directly or indirectly in the manifold initiatives in data-driven technologies

currently unfolding in healthcare - discursively conduct such ethics work offers a valuable addition to literature on ethical dimensions of Big Data and algorithms (Hoeyer et al., 2017). We conceptualize this with the notion of 'ethical framing', which we define as the discursive work conducted by actors to characterize the ethical dimensions of data-driven healthcare technologies, thereby implicitly or explicitly assigning tasks and roles for stakeholders and shaping the promissory future of data-driven healthcare technologies.

Conceptually, this chapter seeks to bring together the hitherto separated themes of promissory futures and an emic approach to ethics as discursive work in the context of data-driven healthcare technologies. Empirically, the chapter is based on 137 interviews and document analysis conducted within an international research project, in which we compared the regulatory and governance dimensions of health-related Big Data in eight European countries. While work on sociotechnical futures in STS has been instrumental in explicating their inherent normativities and performative dimensions (Borup et al., 2006; Brown and Michael, 2003; Van Lente, 2012; Jasanoff and Kim, 2009, 2015; Felt, 2015), thus destabilizing all-too-sticky narratives of progress and revolution (cf. Jerak-Zuiderent, 2015; Stevens et al., 2018), this work has focused less explicitly on the role of ethical framing in shaping promissory futures.

The chapter addresses three empirical questions: *(1) Which ethical frames are enacted by key actors in data-driven healthcare technologies? (2) What consequences follow from these ethical frames in terms of the roles and responsibilities these actors envisage for themselves and others? (3) How do these ethical frames shape the promissory future of data-driven healthcare technologies?*

In the next section, we further contextualize the theoretical traditions the chapter builds on, linking literature on promissory futures, ethics of data-driven technologies, and ethical framing as discursive work. Next, we discuss the methods and introduce the international research program our analysis is based on. In the results, we outline three widely recurring types of ethical framing and show their effects on roles and responsibilities attributed to different actors. In the discussion, we outline the added value of bringing the distinct bodies of literature on promissory futures and ethical framing together in the context of data-driven healthcare technologies and outline themes for new research.

## Theoretical framework

### New technologies and their promissory futures

The future has been an important trope in STS research. From various perspectives, authors have analyzed how visions of the future of technology become articulated in imaginaries and narratives (De Wilde, 2000; cf. Pollock and Williams, 2010).

The sub-field of STS that has engaged most explicitly with promissory futures is the sociology of expectations. The sociology of expectations literature focuses on the informal production and circulation of expectations in science and technology (Borup et al., 2006). Key characteristics of expectations include their performativity, temporal variability and socio-spatial variability (Van Lente, 2012). Expectations can legitimize investments, set courses of action, steer the options to be explored by researchers and technology developers, and coordinate activities and role divisions of research and industry networks (Van Lente, 2012). This work emphasizes that expectations harbor values, norms and provoke role divisions.

To analyze promissory futures, the notion of sociotechnical imaginaries (Jasanoff and Kim, 2009, 2015) proved to be productive. Sociotechnical imaginaries are defined as “collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology” (Jasanoff, 2015: 4). Originally envisaged as closely tied to nation states, later work emphasized that sociotechnical imaginaries can also be articulated by other organized groups and communities (Jasanoff, 2015), can become regionally ‘translated’ in distinctive ways (Levenda et al., 2019), and may acquire particular meanings in specific technopolitical cultures (Felt, 2015). In the context of healthcare, several authors have noted that on a European level, broadly similar discourses on the promissory future of data-driven healthcare circulate (Rieder, 2018; Stevens et al., 2018; Stevens et al., forthcoming).

Taken together, these bodies of work offer useful analytical anchor points to understand the functions and effects of promissory futures. The sociology of expectations literature emphasizes the performative and political dimensions of the informal expectations that circulate through promissory futures, while the literature on sociotechnical imaginaries emphasizes how expectations become durable and stabilized in narratives that can be circulating on an international level, country-specific, regionally bound or tied to particular groups.

## Ethics of data-driven technologies

Ethical concerns regarding data-driven technologies have received a relatively large amount of attention in the fields of STS, CDS and CAS. Several review articles have identified key areas of concern and have sought to map the debate for particular aspects of data-driven technologies (e.g. algorithmic decision-making) and in particular contexts (e.g. biomedicine) (Mittelstadt and Floridi, 2016; Mittelstadt et al., 2016; Mittelstadt, 2019). Work on the ethics of data-driven technologies emphasized not only the wide range of ethical dilemmas that should be considered, but also focused on further theorization (leading to important discussions on how data-driven technologies can be conceptualized) and problematization (highlighting the misalignments with traditional ethical principles and theories). Yet, there is much less focus on how ethical arguments are used as part of discursive work by various actors engaged in the development, implementation and governance of data-driven healthcare technologies (cf. Cool, 2019; Hoeyer et al., 2017).

For the aim of this chapter, we address several insights that have centered on three ways in which data-driven technologies are perceived to problematize traditional ethical notions and theories (cf. Verbeek, 2014). First, many articles argue that ethical aspects should not be viewed as external to data-driven technologies but emphasize the intertwining of technologies and ethical dimensions. Traditionally, such work has focused on the 'hidden normativities' in data-driven technologies (cf. Ziewitz (2016) for a critique on this popularization of what he calls the 'algorithmic drama'). More recently, the focus has been on how this intertwining co-constitutes new relations between human and non-human actors in diverse practices (Ananny, 2016; Lee et al., 2019; Lee and Larsen, 2019).

Second, many articles focus on the 'unprecedented' character of ethical dilemmas, thus questioning the ability of traditional ethical principles and procedures to adequately deal with 'new' dilemmas. For instance, Metcalf and Crawford (2016: 2) argue how data-driven technology "fundamentally changes our understanding of research data to be (at least in theory) infinitely connectable, indefinitely repurposable, continuously updatable and easily removed from the context of collection". Such characteristics are fundamentally at odds with the orientation of most institutional review boards and ethics committees, which focus on individual risks (Zook et al., 2017).

Third, recent work also developed critiques on the limited usability of classical ethical theories in the context of data-driven technologies. A core focus of this work has been on explicating misalignments between data-driven technologies and traditional ethical principles that emphasize moral agency and come with a strong set of assumptions about individualism and free will (Zwitter, 2014; Prainsack, 2018). Such assumptions become especially problematic in the light of data-driven technologies, where the input used for



decision-making is opaque, effects are hard to oversee and unequally distributed, and responsibility, trust, and agency become dispersed.

Absent from this literature is an exploration of how ethical arguments become tied to the promissory future of data-driven healthcare technologies while promissory rhetorics also shape the way ethics is being done and understood. We therefore argue for the analytical productivity of bringing the dimensions of ethical arguments and promissory futures together via the notion of ethical framing.

### **Framing as discursive work**

The notion of 'frame' has a long history in STS (Orlikowski and Gash, 1994; Bijker, 2006) and in interpretive policy studies (IPS). In the latter, the notion originates in both the symbolic interactionist tradition of Goffman (1974) and the public policy analysis developed by Rein and Schön (1993). Goffman used the notion of 'frame' to understand how people implicitly 'define the situation' in interpersonal encounters as they negotiate the meaning of their interactions. The policy analytic approach of Rein and Schön originally focuses on frame analysis as a methodology for the analysis of policy controversies. Through frames, policy actors can highlight certain features of a situation, ignore or select out other features, and bind the highlighted features together into a coherent and comprehensible pattern (Van Hulst and Yanow, 2016). The act of framing then highlights the interactive processes through which frames are constructed through naming, selecting, storytelling, sensemaking and categorizing (*ibid*).

These two scholarship branches (STS and IPS) highlight the discursive dimensions of framing by placing their analytical focus largely on the shared interpretation of the (technological) object or (policy) issue at stake. Framing however, also generates particular tasks and role divisions for various actors, and through framing, roles and responsibilities become distributed and defined in particular ways. Recent ethnographic work in organization studies has sought to explicate how framing as a discursive activity contributes to role divisions, tasks and responsibilities of different actors (Oldenhof et al., 2016; Berghout et al., 2019). Specific to ethics, this point is reinforced by taking into account Wainwright et al.'s (2006) study of ethical boundary work that highlights how scientists conduct ethical boundary work as a process of social demarcation.

To summarize, discursive approaches in IPS and organization studies are instrumental to better understand the specific effects of framing – as a discursive activity - on the interpretation and representation of the technological object or policy issue at stake, and the distribution of specific roles and responsibilities for the various actors involved.

## Methods

The data on which this chapter is based is part of a broader project to investigate the regulatory and governance dimensions of health-related Big Data <sup>1</sup> in a European context. We have conducted 137 semi-structured expert interviews in eight European countries: Austria, France, Germany, Ireland, the Netherlands, Spain, Sweden, and the United Kingdom. Respondents were identified via the network of the project consortium members <sup>2</sup> and via desk research. We selected respondents amongst the categories: (1) healthcare professionals and managers involved in Big Data-pilots that are part of the project consortium; (2) ethical and legal experts who are knowledgeable about the key discussions in their country, (3) technology developers and data scientists who are involved with the technical dimensions of Big Data analyses; (4) representatives of patient and professional associations; (5) visible actors in the public debate in order to capture public perspectives on Big Data; (6) policymakers and additional policy experts who can provide input on key governance strategies and challenges.

The goal of the semi-structured interviews was to develop insights into governance approaches, regulatory challenges, ethical dimensions and societal debates about Big Data, resulting in one infographic per country. <sup>3</sup> Upfront a predefined topic list was made including (1) policy goals and strategies; (2) the main regulations for health-related Big Data and how such regulations are perceived; (3) ethical questions and debates that are generated by Big Data; (4) reflection on how social and cultural aspects are perceived to influence the practices and regulation of Big Data. The large majority of interviews was conducted face to face, with interviews conducted via video conferences or phone in exceptional cases. In all cases, permission to record the interview was obtained. All interview transcripts are stored in a secure server (Workspace).

A supplementary document analysis was conducted for each country to become familiar with the different health systems, public discussions and policy strategies of health-related Big Data. Interview data were triangulated with policy documents, news articles, scientific papers, presentations and grey literature provided by the respondents.

In the analysis of the material, we used a method of constant comparison (Strauss and Corbin, 1990) to facilitate peer review between team members, increase mutual understanding and facilitate in-depth comparison. We organized several full-day meetings to present our initial analysis and reflect on similarities and differences between countries. The theme of ethical framing emerged inductively as one interesting aspect in the data. We were triggered by several remarks of respondents who explicitly referred to what they viewed as a useful approach to the ethics of Big Data, thereby often ascribing a particular

position to ethics and ethical arguments. We therefore conducted additional abductive analysis (Tavory and Timmermans, 2014), iterating between our empirical material and theoretical work on ethics of Big Data and the variety of STS-literature on sociotechnical futures. In this process, we gradually zoomed in on ways of ethical framing.

We interpreted 'ethical' statements broadly in our material, focusing in general terms on what respondents described as things we ought to do or when they referred to things they considered good, valuable, worthwhile or laudable. It was not our intention to define ethics upfront. Instead, we took inspiration from Heeney (2017), who does not attempt to distinguish between the 'real' ethical position and interviewees' interests, but instead states that she is "interested in the discourses interviewees employ and create, and [...] use[s] the interview data to consider what they think ought to be considered in relation to what is and is not ethical in terms of practice" (2017: 11). We recognize with her that interviewees appear not only as 'practicing actors' but also as 'thinkers' who relate their desires for their own practices with the wider normative, scientific, and social implications of what they do (ibid.).

## Results

We structure the results around three widely recurring types of ethical framing: ethics as 'balancing act', the technical 'fix', and ethics as 'collective thought process'. We outline the characteristics of these frames and discuss their implications for the envisaged roles and responsibilities of various actors.

### Framing strategy I: Ethics as 'balancing act'

*"I think that in general [...] we have to develop a balance. [Big Data] is not something where you can have everything: full privacy, full usage of all data and knowledge gain. We have to bring it in balance. First, there is data protection. Here [we] must find the balance between protection for patients' privacy and the possibility to do research with it. [...] It is the task of the policy and lawmakers to find a good balance. So that those that are feeling threatened feel kind of safe, but at the same time, the researchers can do what they have to do. [...] Also ethically, I have to balance it. [...] It is nothing else than finding a balance between risk and benefit" (Interview information scientist, Austria).*

The first type of framing we encountered is ethics as a balancing act. The perceived role for ethics in shaping the promissory future of data-driven healthcare technologies is to

balance. This framing entails discursive work in which different ethical values or principles are outlined in relation to each other, as exemplified in the above quote. Such relations are, then positioned as either being weighed differently in various contexts, as being in tension with each other, or as being opposite ends of a spectrum. Ethical use of data-driven healthcare technologies according to this frame then presupposes the search for an adequate mean between various values, or, in other terms, about finding an appropriate balance. As different values become outlined in relation to each other, the 'ethical' thing to do becomes envisaged as an attempt to find the right balance: between economic values (such as innovation capacity), public values (such as privacy and data protection), professional values (such as delivering good care), scientific values (such as knowledge development, learning), and legal values (such as data integrity). Respondents implicitly assign the task to all stakeholders involved; the quote expresses this role as both a general assignment ("we have to") and a personal obligation ("also I").

The promissory future of data-driven healthcare technologies is presented as a mixed set of benefits and risks that are not neatly aligned. In some variants of this frame, the balancing act is explicitly presented as a precarious and temporary achievement; like a pendulum, the ethical focus within a country is argued to shift from an emphasis on one value to another. The following quote from a professor in health information in Ireland expresses this variant of the balancing frame as he reflects on what he views as a gradual movement from the emphasis on one ethical value to the current overemphasis on another:

*"[In Ireland] there is a fear to the point of paralysis of sharing data because of data protection issues. A lot of it is unfounded, but everyone kind of fears data protection and therefore [many organizations] are saying: no, I am not sharing my data. [...] So the pendulum has swung far too far over to the protection of the patient and I think we are potentially at the point where we are not protecting the patients' care. We are protecting their data, but we may not be protecting their care. So there is a sweet spot in between those two. You know, in the past, I'd say we could have been more on the other end you know: share everything and no control. I think we probably are going too far at the other way as well" (Interview professor health information, Ireland).*

This example shows how a discursive contrast is created between different values, but in this example, they become linked to a (critical) narrative of gradual change over time and overemphasis on the ethical value of data protection at the perceived cost of the ethical value of patient protection.

The frame of the 'balancing act' can also be used as a strategic attempt to reframe the ethical questions that are considered to matter. For instance, various high-level policymakers discursively attempted to reframe the ethical debate towards a recognition that not sharing available data is as ethically troublesome as sharing data:

*"We can't keep hiding behind the ethics. We say that it is not ethical to use people's data, but maybe it is just the other way around. It is not ethical not to use the data. So I think that this is sort of changing and people are very open to donate to healthcare. [...] If we say that it's not ethical to use my data in a wider context, it is always very easy to hide behind that, to say: "no, we can't move in this direction with utilizing data in this way". [...] It's easy to stay where we are right now and say that this is the right ethical framework. But I think from a patient perspective, it's not ethical not to use the data" (Interview policymaker county council, Sweden).*

This quote shows a policymaker implicitly utilizing the frame of ethics as a 'balancing act' as a way of reframing the current debate on data sharing, which according to her focuses too much on values like privacy and data protection. It can also be understood as an attempt to depict the current position of ethicists as 'easy': critically commenting on ethical questions and dilemmas raised by data-driven healthcare technologies is depicted as one-sided and as a way to 'hide' from the discussion about whether it is not equally unethical to create barriers for data sharing, as this is argued to lead to patient risks. This reframing also comes with frequently used metaphors such as 'data donation', highlighting 'collective' principles and discursively linking ethics of data sharing to notions of 'good citizenship' (i.e. as something that 'responsible' citizens do) (cf. Puschmann and Burgess, 2014). Such collective principles again tie into the framing of ethics as a 'balancing act' as they highlight that the emphasis is currently placed too much on individual principles.

An important implication of framing the role of ethics as a balancing act is that values are perceived as comparable. This way of framing thus assumes that values can be 'weighed' against each other in a utilitarian cost-benefit type of analysis. As such, this way of framing offers less space for values that are considered to be incompatible. The search for the adequate balance mostly implies that more of one type of value leads to less of the other type of value.

This way of framing also has consequences for roles and responsibilities implicitly aligned to ethicists and other actors. The role of ethics implicit here is one that is well-demarcated, but simultaneously minor, as ethics becomes primarily about finding and maintaining the right balance between ethical values in particular use contexts of data-driven healthcare

technologies. The quotes make clear that balancing is not necessarily tied to ethical experts: policymakers and, indeed, any actor can engage in ‘ethical reasoning’ of this kind. While this act of framing can be perceived as an attempt at democratization, it also obfuscates aspects that various scholars would describe as important for ethicists in their role as experts. For one, this framing neglects the role of ethical experts as actors whose value lies in their ability to disrupt all too optimistic or instrumental narratives (cf. Swierstra, 2018). Another aspect of ethics that becomes marginalized in this way of framing is the attention for ethical dilemmas that do not neatly ‘fit’ in the implicit cost-benefit structure of the frame, and for which again many scholars would argue a particular kind of ‘moral’ or ‘ethical expertise’ is warranted (Grunwald, 2004; Prialux et al., 2016). Thus, this way of framing comes with an implicit set of tasks and role divisions that highlights the role of ethics as ‘ethical reasoning’ any actor can engage in, while neglecting a more ‘fundamental’ role of ethics as ‘disruptor’.

## **Framing strategy II: The technical ‘fix’**

*“I guess one of the difficulties we have at the moment is our ethics committee here. Our institutional review board is still grappling with GDPR regulation [...]. So, while they are still grappling with that, there is a little bit of stagnation with the ethics process. And that can be [troublesome], but none of that is insurmountable. [...] I guess everyone is figuring it out together you know, which is fine. It’s a bit uncharted territory [...]. But once we can demonstrate that we are compliant in anonymizing everything, not making it linkable back to the patient, then I don’t foresee a problem. And nor did our data protection officer” (Interview physician, Ireland).*

The second type of framing we encountered is the frame of the ‘technical fix’. The perceived role for ethics in shaping the promissory future of data-driven healthcare technologies is to solve. The above quote from an Irish physician, working on a Big Data project to reuse hospital data for analytical purposes, is illustrative for this type of framing. This framing entails discursive work in which the ethical problems with data-driven healthcare technologies are recognized, but simultaneously presented as ultimately solvable through technical or instrumental means. Thus, while many actors recognize ethical risks of re-identification and lack of informed consent, these ethical risks are not viewed as insurmountable. Rather, given the right instrumental and technical solutions, the idea is that ethical risks can be contained and dealt with adequately via existing institutional structures such as ethical review boards.

Perceived solutions frequently envisaged within this type of framing cluster around three approaches. First, better technical solutions. Some respondents focused, for example, on techniques for anonymization or pseudonymization as they were perceived as sufficient safeguards for the ethical use of health data. Second, better procedures that clarify the stakes or seek to navigate around the most problematic ethical aspects. For example, by asking 'broad' consent forms for research areas as a 'solution' for the ethical problem of repeatedly asking for individual consent. Lastly, developing proper legislative frameworks and principles. This can be seen, for instance, in the way discussions about 'proper' (i.e. 'ethically acceptable') use of health data congregated around the narrow question of whether practices aligned with the new European General Data Protection Regulation (GDPR) or not.

As with the framing of ethics as 'balancing act', the framing of the 'technical fix' also comes with various implications. First and foremost, this way of framing comes with the normalization (and with this the 'deflation') of ethical dilemmas. Ethical issues are implicitly perceived as 'fixable' with additional procedures, guidelines, instruments or legislation. As such, this way of framing can result in 'ticking-the-box-ethics', even though experts question the saliency of such solutions (cf. Zook et al. (2017)'s final 'rule'). Second, the belief in a technical fix in this way of framing backgrounds ethical processes of joint deliberation and compromising. Third, the role of ethics becomes narrowed to 'fighting symptoms', as this way of framing emphasizes ethical aspects that are relatively easily 'doable', instead of the more complex ethical issues that are often raised in more critical literature on data-driven technologies (boyd and Crawford, 2012; Mittelstadt and Floridi, 2016; Mittelstadt, 2019; Grote and Berens, 2020).

As with the previous type of framing, the frame of the technical fix also comes with consequences for the roles and responsibilities assigned to ethicists and other actors. Similar to the framing of ethics as a balancing act, this type of framing also implies a well-demarcated and relatively minor role for ethics. In essence, all three variants implicitly render ethical expertise subordinate to other domains or at least render ethical issues and dilemmas 'solvable' through other domains: whether by technicians (who can develop the 'right' solutions), by policymakers and health managers (who are positioned to develop such 'better procedures'), or by lawmakers (who can develop 'proper legislative frameworks'). Such implicit task distributions come with risks as well. A potential risk exists in equalizing the development of the 'right' procedure as such with the idea of 'ethical use' of data. While the development of transparent procedures is undoubtedly valuable in terms of specifying ethical criteria for access and restriction of sensitive health data, the procedure as such is unlikely to ameliorate the broader ethical discussions, for example, about forms of opacity, pernicious feedback loops that reinforce effects of inaccurate data, bias, or

discrimination, and the ways in which professionals may become narrowly embedded in 'epistemic niches' that constrain or undermine the practical wisdom necessary to make moral decisions (Hayes et al., 2020).

### **Framing strategy III: Ethics as 'collective thought process'**

*"Ethics is an eminently collective, plural concept. [Therefore], [...] the CNIL could not lay claim to any sort of monopoly over ethical discussions on digital technology. On such a vast and cross-cutting subject, on no account should these be held behind closed doors. [...] It was in this mindset that the CNIL set a collective approach in motion for several months overseeing a public debate with the help of partners from various sectorial fields (health, justice...). In this respect, ethics is just as much about the process itself as it is about the outcome" (Commission Nationale de l'Informatique et des Libertés, 2017: 3-4).*

The third type of framing we encountered, is the frame of ethics as a collective thought process. The perceived role for ethics in shaping the promissory future of data-driven healthcare technologies is to deliberate. The above quote from the Commission Nationale de l'Informatique et des Libertés (CNIL), the French national data protection agency, is illustrative for this type of framing. This framing entails discursive work in which ethics becomes viewed as an open-ended process of joint deliberation. In the example above, the CNIL has organized a series of public debates about the development of new digital technologies. It has explicitly positioned ethics as the "crafting [of] a collective and pluralist ethical thought process" (CNIL, 2017: 8). Key underlying ideas in this way of framing are notions of deliberation and continued reflection, which are perceived as crucial in engaging with the ethical dimensions of data-driven (health) technologies in a 'good' way. Through its emphasis on the processual dimensions of ethics as a form of deliberation, this frame also recognizes the open-ended character of ethics: technological developments are perceived as fundamentally uncertain and thus likely to require continued attention and revision.

Aligned with this way of framing is a critique on a narrow expert focus. A policy approach often seen is to 'outsource' ethical discussions to expert advisory organizations and committees. In Austria, for example, discussions about new data-driven healthcare technologies are mostly organized in terms of expert deliberation. A patient representative working at an advocacy organization enacts the frame of ethics as a collective thought process to question this narrow focus on experts:



*“So, the debate [on data-driven healthcare technologies] takes place in a very small group, so it does not take place in the general population [...]. But then it is immediately...when it is no longer an urgent topic, it breaks down relatively quickly. So, the general population doesn't really take part in the debate. These are debates that actually run among the experts” (Interview patient representative, Austria).*

In this way of ethical framing, the tendency to limit ethical debates to experts is explicitly problematized. This line of reasoning can also be seen in the argumentation of the CNIL, as the institute outlines how such limitations may lead not only to public mistrust and suspicion, but also because expert perspectives are insufficient given the complexities and uncertainties in the evolution of digital technologies (CNIL, 2017: 4).

Again, this way of framing comes with a set of implications. The implications of this frame are primarily that it asks for a democratization of ethics, with the aim to include not only ‘ethical experts’, but to open up the realm of ethics to many stakeholders, including laypersons affected by new technological developments. Often underlying this way of framing is an implicit rationale that new technologies will affect many citizens and that this entitles them to become involved in ethical deliberations. As such, ethics and ‘citizenship’ become discursively entwined in this way of framing. Another implication of this way of framing is that it explicitly allows for – and even seeks to organize – forms of multiplicity and diversity.

As with the previous ways of framing, the frame of ethics as a collective thought process comes with consequences for the roles and responsibilities assigned to ethicists and other actors. Most notably, it comes with a broadening of responsibility to a much more varied set of actors. As ethics becomes a public concern that is not limited to a set of expert actors, nor something that can be fixed procedurally or technically, anyone can – and one might even say ‘should’ – become involved in ethical reasoning. In this sense, the frame explicitly distances itself from a perspective on ethics as belonging to a specialist domain. Simultaneously, the way this frame is enacted can also affirm organizations in taking up a key position. The example of the CNIL shows how this agency positions itself as the ‘custodian’ of key ethical principles and as the organization most suited to organize this process of inclusive deliberation. As the CNIL decides on the core ethical principles, their role becomes one of facilitating debate and crafting key messages to policymakers (which involves a lot of translation efforts; not only summarizing the debate, but also selecting the most important outcomes and tailoring these to policy recommendations). Next to setting the ‘rules of the game’ and translating outcomes, one could argue that this role of ‘custodian’ also involves responsibilities of ensuring ‘voice and choice’ (who is included

and under what terms?) and creating favorable conditions to prevent subtle exclusion strategies (cf. Oldenhof and Wehrens, 2018).

## Discussion

In this chapter, we have sought to bring together STS-literature on promissory futures, ethics literature on data-driven technologies, and policy literature on framing as discursive work in order to study how ethical arguments become tied to the promissory future of data-driven healthcare technologies. This chapter argued for more explicit attention to the role of ethical framing in shaping the promissory future of data-driven technologies. We have identified three widely recurring types of ethical framing: ethics as a balancing act, the technical fix, and ethics as collective thought process. These types of framing come with different perceived roles for ethics in the promissory future of data-driven healthcare technologies: 'to balance', 'to solve', and 'to deliberate'.

With our analysis, we aim to contribute to both ethical literature about data-driven technologies and STS literature on promissory futures. First, we argue that more explicit attention for ethics as a discursive act offers a relevant new perspective to existing ethical literature on data-driven technologies. Existing work has predominantly focused on issues of agenda-setting (making visible ethical dilemmas), theorization (leading to new conceptualizations of data-driven technologies) and problematization (highlighting the misalignments with traditional ethical concepts and principles) (boyd et al., 2014; Zwitter, 2014; Metcalf and Crawford, 2016; Mittelstadt and Floridi, 2016; Zook et al., 2017; Lee and Larsen, 2019; Mittelstadt, 2019). We argue that there has been less focus on how ethical arguments are used as part of discursive work by practitioners engaged in the development, implementation and governance of data-driven healthcare technologies (cf. Cool, 2019; Hoeyer et al., 2017). Our analysis shows how such discursive work, which we have conceptualized through the notion of 'ethical framing', attempts to shape the promissory future of data-driven healthcare technologies in various ways.

In line with insights from the sociology of expectations (Borup et al., 2006; Van Lente, 2012), acts of ethical framing appear to come with similar performative effects: they generate particular types of concerns, they mobilize different people, and they steer the agenda in particular ways. Framing ethics as a balancing act generates as main concern for data-driven healthcare technologies that they come with a mixed set of benefits and risks that need to be carefully aligned. This frame mobilizes policymakers and other stakeholders as rational actors who should engage in utilitarian decision-making. The balancing frame moreover steers the agenda by focusing on finding an 'optimum' between this

contrasting set of benefits and risks. Framing ethics in terms of a technical fix generates different concerns, such as the lack of clear instruments, guidelines and legislation. This frame mobilizes legal experts, policymakers and technicians as key actors and steers the agenda to the provision of such legislation and the development of clear guidelines and instrumental solutions in order to 'solve' remaining ethical concerns. Framing ethics as a collective thought process highlights yet other concerns as this frame stresses the fundamental uncertainty and unpredictability of technological developments. This frame mobilizes all citizens to become involved in deliberative processes and steers the agenda towards such forms of deliberation as crucial to enhance the perceived legitimacy of data-driven healthcare technologies.

Attention to the role of ethics as a discursive act thus offers a valuable additional perspective to ethical literature about data-driven technologies because it places analytical attention towards the effects of framing by various actors. The focus on ethics as discursive practice also leads to new questions, for instance about which actors utilize what kind of ethical 'vocabulary' and what this renders visible and invisible, about how we should evaluate various acts of ethical framing (e.g. under which conditions and under what circumstances do we consider particular acts of framing to be justified or not), and about what types of framing become dominant at which moment in time, and to what effects.

A second body of literature this chapter contributes to, is the STS literature on promissory futures. While the normative dimensions of such promissory futures are widely recognized (Borup et al., 2006; Van Lente, 2012; Jasanoff and Kim, 2009, 2015; De Wilde, 2000), we argue that the focus on ethical framing adds a valuable new dimension to our understanding of these normative dimensions. Our analysis shows how strategies of ethical framing further interact with the performative and political aspects of promissory futures, generating particular concerns and steering the agenda in particular ways. Our study moreover suggests that, in addition to earlier observations that broadly similar discourses on the promissory future of data-driven healthcare circulate in different countries (Rieder, 2018; Stevens et al., 2018; Stevens et al., forthcoming), we can similarly recognize distinctive types of ethical framing that are prevalent irrespective of country-specific or regionally bound traditions and data practices (cf. Tupasela et al., 2020; Felt 2015).

More attention to acts of ethical framing in shaping the promissory futures of new technologies is also likely to expand our analytical focus towards interesting new domains. For one, longitudinal studies could highlight whether and how different actor coalitions are involved in 'framing struggles' over time and how this may affect the perceived cultural legitimacy of data-driven healthcare technologies (cf. Geels and Verhees, 2011). Second, it would be insightful to extrapolate our findings to different domains (e.g. policing,

media, law) as such fields may come with different promissory futures and alternative discursive acts of framing (e.g. Asaro, 2019; Egbert, 2018). Relatedly, our exploration of the relations between promissory futures and acts of ethical framing could be extended to other potentially 'disruptive' future technologies, as these acts of framing are likely to play an important role in the development of such technologies as well. Third, comparative ethnographic studies of different data-driven initiatives and practices could focus on ethical framing 'in action'. There is still much to be learned about how ethical questions are navigated in medical practice, how grounded considerations are made and justified, and how data scientists, technologists and medical practitioners negotiate norms and values in concrete initiatives (cf. Stevens et al., 2020). A focus on ethics as discursive work offers a valuable addition to such ethnographic studies as it can highlight how ethical questions are framed and what roles and responsibilities such frames imply for different actors.

To sum up, this chapter argued for more explicit attention to the role of ethical framing in shaping the promissory future of data-driven healthcare technologies. If readers recognize, as is our conviction, that the prevalent ways of framing the ethical dimensions of data-driven healthcare technologies has wider saliency, we hope to have inspired others to focus on the ways ethical arguments become entwined with the promises of new technologies in other domains as well.

## Footnotes

<sup>1</sup> We utilized a broad interpretation of Big Data, which moves beyond specific technologies for data aggregation and analysis. This broad interpretation is in line with the sociotechnical approach developed in boyd and Crawford (2012). We summarized it here under the term 'data-driven healthcare technologies'.

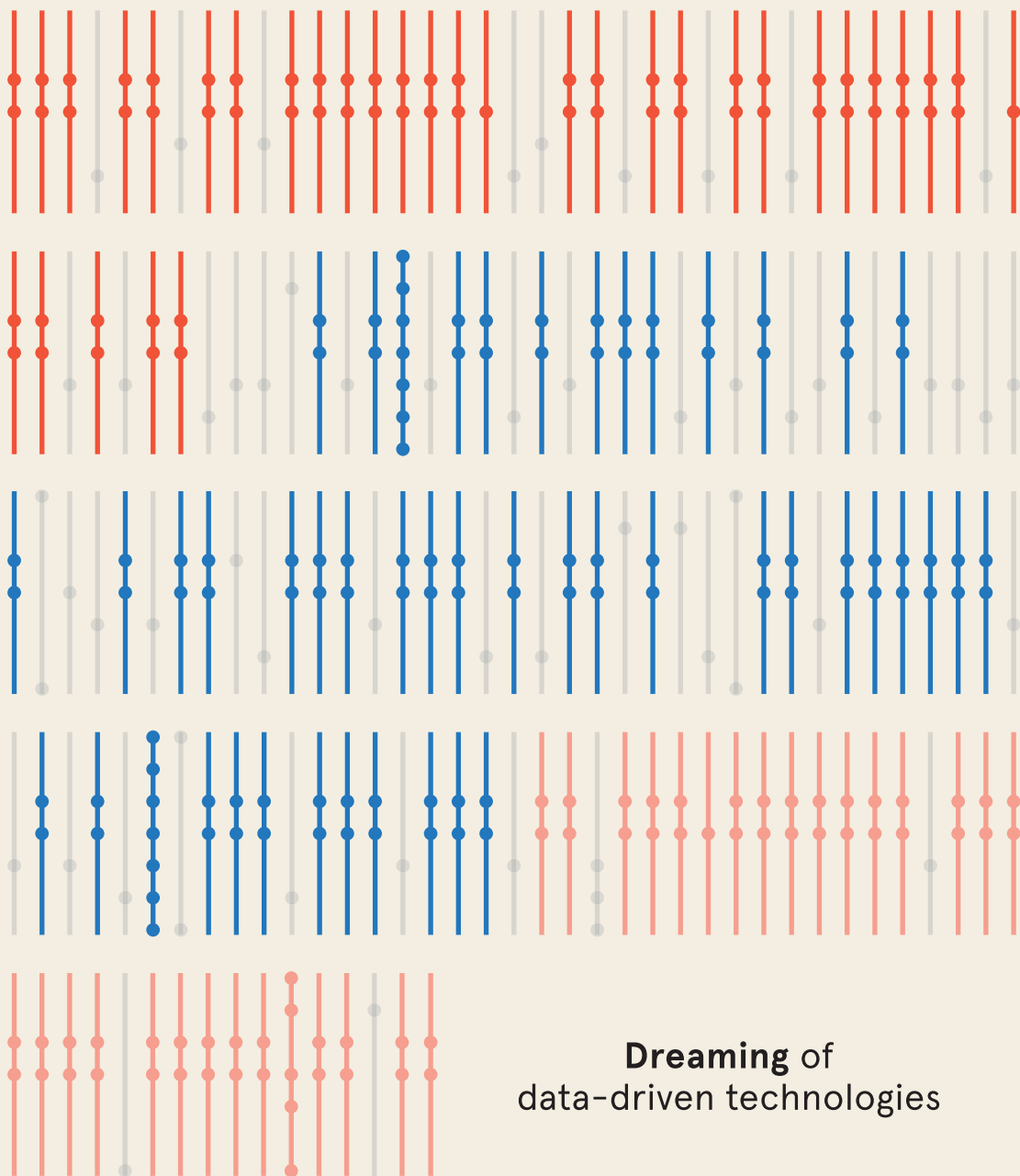
<sup>2</sup> See <https://www.bigmedilytics.eu/> for more information about the consortium.

<sup>3</sup> See [https://www.bigmedilytics.eu/wp-content/uploads/2019/01/D5.2\\_Country-specific-infographics\\_v1.0-including-infographics.pdf](https://www.bigmedilytics.eu/wp-content/uploads/2019/01/D5.2_Country-specific-infographics_v1.0-including-infographics.pdf)



Under review at Big Data & Society as:  
Stevens M, Wehrens R and de Bont A – Data-driven dreams:  
personal sensemaking of data promises for new healthcare information.

# Chapter 4







## Introduction

The 'future of healthcare will be data-driven' is a common imaginary associated with Big Data, data science and related techniques (e.g. Mayer-Schönberger and Cukier, 2014; Stevens et al., 2018). We interviewed 137 professionals from eight European countries involved in a European Union project on data-driven initiatives in healthcare. All our respondents recognized the grand promises of data, shared their own stories with us and explained why they decided to invest time, energy and resources into realizing data-driven healthcare. We argue that personal sensemaking is understudied in CDS and related fields, whereas it is valuable for understanding the persistence and recurrence of data-driven promises. In our study, we set out to explore the sensemaking of diverse professionals with the metaphor of dreams.

For years, the healthcare sector has been subjected to a rhetoric that promises that data-driven techniques will lead to an awe-inspiring healthcare revolution, gloriously improved healthcare planning and increased quality of healthcare delivery (Ames, 2018; boyd and Crawford, 2012; Mayer-Schönberger and Cukier, 2014; McAfee and Brynjolfsson, 2012). While specific terms go in and out of fashion, the underlying assumption stays the same. Technical developments have made it possible to collect all sorts of data about our lives, bodies and health. These data can now be stored, combined and analyzed in various ways leading to new information, insights and certainties that can be used to improve healthcare in many ways (Mayer-Schönberger and Cukier, 2014).

Scholars in CDS, STS and related fields argue that we should look beyond the hyped nature of promises, but focus instead on the dynamics of the promises to study what they produce and neglect (e.g. Ames, 2018; Smits, 2006; De Wilde, 2000). These scholars have developed ways of studying and theorizing on promissory rhetoric, such as critically analyzing future-oriented discourses (Smits, 2006; Stevens et al., 2018), focusing on the performative effects of buzzwords (Bensaude Vincent, 2015; Hoeyer, 2019; Penkler et al., 2019) and expectations (Borup et al., 2006; Van Lente, 1993; 2012) and exploring the analytically productive notion of sociotechnical imaginaries (Jasanoff and Kim, 2015). This work continually emphasizes that promissory rhetoric is particular to a specific time and place, hides all sorts of normativities and is a huge driver for change because of its performative effects.

While this work does offer valuable insights, we argue that it is as important to pay attention to the diverse professionals who work daily on making data-driven healthcare a reality. These people are *motivated* by the overarching promises in media, scholarly journals and policy documents, for instance. They attach *value* to these promises, and take *consider-*

ations into account before investing time, energy and resources into making data-driven healthcare come true. Professionals do not naively believe all the amazing prospects for the future, nor will they work on nonsensical or unfeasible projects. Their motivations, values and considerations – professional sensemaking – offer valuable insights into the nature of data-driven promises.

Data-driven promises are omnipresent in European healthcare. Indeed, the European Commission actively stimulates data-driven initiatives (Rieder, 2018). Emphasizing that Big Data are crucial to maintaining EU competitiveness (EC, 2014; Rieder, 2018) in 2017 the European Commission decided to fund a large, multinational public-private partnership in healthcare in an effort to “transform healthcare through Big Data” (BigMedilitics, 2020). This led to 12 data-driven healthcare initiatives, concentrated in eight European countries, intended to demonstrate the added value of intensive use of data technology. We were able to conduct interviews with diverse professionals working on these data-driven initiatives. The interviews helped us answer the following research question: *how do professionals make sense of the promissory rhetoric of data-driven information in the healthcare sector in eight European countries?*

We introduce the metaphor of dreams to present insights into the personal sensemaking of data-driven promises and answer this question. Although the concept of dreams is often used nowadays, it is underdeveloped in CDS as illustrated by the many publications in the journal of Big Data & Society that mention or refer to dreams but do not theoretically develop the term. “Dreams” are used as an adjective (e.g. dream house) (Steeves, 2020), to indicate particular visions of the future (Gillespie, 2020; Koed Madsen, 2018; Rexhepi, 2016; Wadmann and Hoeyer, 2018; Wilmott, 2016), to mark naïve portrayals of technological developments in contrast to reality (Ames, 2018; Aradau and Blanke, 2015; Chalmers and Edwards, 2017; Seaver, 2017; Törnberg and Thörnberg, 2018; Venturini et al., 2017) or seen as highly personal just like our thoughts and memory (and, therefore, should be protected against invasion of privacy and overly dominant technological approaches) (Andrejevic, 2020; Barnes and Wilson, 2014; Thrift, 2014).

In this wide yet unspecified range of usage, we see two recurring characteristics that make the metaphor of dreams useful for our purpose. First, we see that the term is often linked to positive portrayals of a future that is different from the current situation. Second, dreams are often seen as specific and highly personal. We argue that this combination (both highly personal and about envisioning different worlds) makes dreams a valuable metaphor to obtain insights into the personal sensemaking of data-driven promises.

This chapter is relevant for CDS scholars in that it develops the metaphor of “dreams” to understand the personal motivations, values and considerations that play a role in data-driven initiatives. Thereby, we answer the call by Jasanoff and Kim (2015) and Rieder (2018) to gain a better understanding of Big Data imaginary, its origins, values and inscribed politics. The insights from this study may help ensure that data-driven technologies become and stay responsive to societal needs. They may help stakeholders decide on policy for the future, encouraging explicit connection with professionals in the field and understanding their motivations, values and considerations with regards to new techniques.

In what follows, we sketch the theoretical background of the chapter and our methodology. We will elaborate on concepts used to study promissory rhetoric and explain the added value of the metaphor of dreams. In the results section we discuss how professionals link data-driven promises to longstanding issues in healthcare and explain the roles of the professionals’ motivations, values and considerations. Finally, in the discussion, we reflect on the metaphor of dreams and consider what paying more attention to personal sensemaking adds to CDS literature.

## Theoretical framework

### Promissory rhetoric

Scholars in STS, CDS and related fields have developed various ways of studying and theorizing on the future. This work emphasizes that promissory rhetoric are particular to a specific time and place, hide all sorts of norms and are simultaneously a huge driver for change because of their performative effects. It teaches that we should be critical of the celebratory rhetoric that surrounds data-driven technologies and drives their implementation by pointing out its unrealistic nature, opening up (implicit) assumptions and explicating changing social dynamics (e.g. Ames, 2018; Stevens et al., 2018). This work is helpful in exploring alternative, more diverse portrayals of the future and articulating counter stories (Konrad et al., 2017). We will illustrate this by elaborating on the value of future-oriented discourses, buzzwords, expectations and sociotechnical imaginaries that have been important in this expanding body of work.

First, scholars have focused on *future-oriented discourses* that arise with new technologies (e.g. Achterhuis, 2002; Ames, 2018; De Wilde, 2000; Smits, 2006). Future-oriented discourses are often overly positive or negative (e.g. Smits, 2006) as is recognizable in the promissory rhetoric that surrounds data-driven technologies (e.g. Ames, 2018). Through discourse this body of work highlights how interpretive social repertoires are formed and given meaning. Discourses play an important role in shaping conceptions of technological

abilities and determining what role the technologies may play in society. In analyzing these discourses and related metaphors, the scholars stimulate more nuanced discourse on the social implications and consequences of technology (e.g. Stevens et al., 2018).

Second, recent work focuses on particular discursive constructs, namely *buzzwords* (Bensaude Vincent, 2015; Hoeyer, 2019; Penkler et al., 2019). These scholars look at concepts such as Big Data, data science and algorithms that can be characterized by a lack of conceptual clarity. Instead of defining and specifying these terms, they argue that we should look closely at the dynamics of buzzwords and what they produce. The scholars argue that it is precisely the inconsistency and vagueness of buzzwords that allow them to act as boundary objects that can bring actors from different social worlds together and coordinate actions (Penkler et al., 2019). Buzzwords steer fields in particular directions, set attractive goals and create matters of concern (Bensaude Vincent, 2015). Thereby, they create new collectives (Penkler et al., 2019) and can help to overcome value conflicts (Bensaude Vincent, 2015).

Third, work on *expectations* traces how both formal and informal expectations shape technologies and technological fields (e.g. Borup et al., 2006; Konrad et al., 2017; Van Lente, 1993; 2012). Actors and institutions expect specific futures. These scholars use expectations to analyze the anticipatory practices of these actors and institutions. This body of work builds on concrete empirical cases to show how expectations mobilize actors, but also legitimate and justify investments, set courses of action, and steer the options to be explored by researchers and technology developers (Konrad et al., 2017; Van Lente, 2012).

Lastly, *sociotechnical imaginaries* have gained attention as an approach that highlights the temporal, situatedness and political dimensions of future-oriented thinking (Jasanoff and Kim 2009; 2015). Sociotechnical imaginaries are defined as “collectively held, institutionally stabilized and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of advances in science and technology” (Jasanoff, 2015: 321). Imaginaries are used to explore, mostly on the national level, the ways collective imaginaries legitimate certain policies while foreclosing others. This has resulted in studies that compare imaginaries and case studies that explain how particular imaginaries arise, circulate, gain traction and are sometimes contested in particular countries or social groups (Jasanoff, 2015).

We argue that for a critical analysis that strives to ensure that data-driven technologies are responsive to societal needs, it is important to pay attention to the personal and collective processes of various actors in sensemaking. The literature described above teaches that promissory rhetoric profoundly affects practice. However, neglected within this body of

literature is the role of personal sensemaking of various actors in their daily work. The literature seems to assume that people working in the field on these technologies are “mesmerized” by the utopian discourses and imaginaries or “mobilized” through buzzwords or expectations. We argue that analytically this is unsatisfactory. Professionals are not known to naively follow the promissory rhetoric for the future, nor are they willing to participate in projects that are nonsensical or infeasible. They tend to be highly critical of the adoption of new technologies and other innovations if these cannot be embedded usefully in the daily flow of their work (e.g. Berghout, 2020; Stevens et al., 2020).

Current attention to individuals in scholarly work on promissory rhetoric is limited to “key” actors with particular roles in the origination and spreading of promises. For example, Jasanoff talks about “daring individuals” (2015: 324) who play a role in the origin of certain sociotechnical imaginaries in a certain time and place. Similarly, Konrad and colleagues explain that “strategic voicing and dedicated promotional efforts” (2017: 487) of individuals are important for developing expectations.

Alternatively, we focus on the sensemaking of data-driven promises of the more “ordinary” actors. Attending to the personal sensemaking of, in our case, professionals gives additional insight into the motivations (why are people mobilized by the promises?), values (why do they find the promises important?) and considerations (what are the trade-offs that they make?) that professionals consider important. These insights help us understand why certain futures do gain traction while others do not, and are thus useful in formulating and prioritizing different futures.

## The metaphor of dreams

We argue that the metaphor of dreams is valuable in studying the personal sensemaking of grand rhetoric. As declared in the Introduction, scholarly work in CDS frequently mentions “dreams” but rarely develops the term theoretically. In the use of the term, we noticed that dreams are about envisioning other worlds yet, at the same time, are highly personal. This combination makes dreams a valuable metaphor to obtain insights into the personal sensemaking of data-driven promises.

Our use of “dreams” lies close to the term Jasanoff (2015) uses occasionally in her work on sociotechnical imaginaries. She describes how individuals can “dream up new worlds” (Jasanoff 2015: 324) and how some dreams can grow into imaginaries, or as Jasanoff puts it: “Imaginaries by definition are group achievements (...). Individual dreams and aspirations take hold and acquire collective force only when key actors mobilize the resources for making their visions durable” (Jasanoff 2015: 25). However, attention to dreams is largely

lost in the analyses of sociotechnical imaginaries as these studies focus on collectiveness and dynamics of imaginaries.

The role of dreams in the work of Jasanoff (2015) emphasizes the generative nature of imaginaries, yet, we would argue that dreams can have a broader function. Literature from dream scholars (e.g. Barrett, 2020; Bulkeley, 2019; Nielsen, 2020) theorizes that dreams have a complex role in making sense of all sorts of experiences. Dreams have a range of functions, such as replaying fragments of experiences, solving problems and regulating emotions. Recent work describes, for example, how the recent COVID-19 pandemic influences our dreams, how much we dream and how well we remember them, and that these dreams also can reinforce certain perceptions of the pandemic (Barrett, 2020). Based on this work, we argue that the dream is not only valuable for understanding the origin of a promissory rhetoric, but also for understanding the broader sensemaking process of the rhetoric. This means that the metaphor of dreams also can give insights into the growth, circulation and debates of technological promises.

To summarize, while current work on promissory rhetoric is highly useful to gain a better understanding of how shared promises influence the practice of particular technological innovations, this body of work neglects personal sensemaking processes while insight in the personal sensemaking of technological promises is valuable in understanding the potential value and future of, for example, data-driven technologies and why people attach particular values to the developments. We apply the metaphor of dreams in this chapter to gain insight into the underlying motivations, values and considerations that lead people to participate in technological innovation.

## Methods

### Context

As mentioned in the Introduction, the European Commission stimulated the formation of a large multinational public-private partnership in healthcare that consists of 35 consortium parties working in diverse European countries. The goal of this project was to (1) use Big Data technologies to improve productivity in the healthcare sector, (2) ensure long-term affordability of healthcare services and (3) stimulate the adoption of Big Data usage in the sector (BigMedilytics, 2020). As one of the consortium parties, we (the authors) have investigated various social dimensions of data-driven technologies. We focused on eight countries in which the European partnership concentrated their efforts: Austria, France, Germany, Ireland, the Netherlands, Spain, Sweden and the United Kingdom. This study uses the data collected in these countries.

## Data collection

We divided the countries among the members of the research team, basing our approach on a combination of qualitative methods (Wehrens et al., 2020). First, we conducted desk research, analyzing policy documents, news articles, scientific papers and gray literature to become familiar with the different health systems, the public discussions about data-driven technologies and the diverse perspectives and positions of the countries. This document analysis resulted in draft reports about data-driven technologies for each country.

Second, between April and October 2018, we conducted a total of 137 interviews with professionals involved in data-driven technologies. The goal was to form an understanding of the governance, regulatory, ethical and social approaches and challenges regarding data-driven technologies for each country. This meant that we were interested in national policy strategies, rules and regulations, ethical debates and societal discussions that surrounded the technologies. We identified respondents via the members of the 35 consortium parties and our document analyses. We strived for a balance between different professional categories (healthcare professionals, ethical and legal experts, technology developers and data scientists, patient representatives and visible actors in the public debate and policymakers), yet also chose to be pragmatic and inclusive in our approach. The interviews lasted between 30 and 180 minutes.

Most of the interviews were conducted during country visits as this meant that respondents could show us their organizations, their projects and we could quickly build rapport. Some interviews were held online or by phone. The semi-structured interviews were based on a predefined topic list, focusing on several core themes yet allowing room for flexibility and adjustment. At the start of each session, we asked permission to record the interview. In all cases, permission was obtained, and the interviews were recorded and subsequently transcribed ad verbatim.

## Data analysis

As team members, we found it important to work closely together throughout the project. We conducted a thematic analysis of the material, using the method of constant comparison (Corbin and Strauss, 1990; Deville et al., 2016) to facilitate peer review between team members, increase mutual understanding and facilitate in-depth comparison. This meant that throughout the project we organized several full-day team meetings to discuss and present our initial analysis and reflect on similarities and differences between the countries. Meanwhile, we sent each other creative exercises, such as “capture, in an original manner, your most challenging interview” or “select an image that reflects the main surprise you encountered in the documents that you have read about your countries”. These exercises often inspired creative findings and helped us to think beyond the obvious lines of thought.

Lastly, we worked intensively together to analyze the material in various ways to develop infographics that summarize relevant regulations and practices (BigMedilytics, 2018) and other papers that focus on ethical framing and legitimacy (Wehrens et al., 2018; Wehrens et al., forthcoming).

During our collaboration, we noticed not only similar promissory rhetoric in all the European countries, but also similarities in the ways that professionals made sense of the rhetoric. This inductive finding made us decide to focus on personal sensemaking in Europe and not to generalize country-specifics for this chapter. What followed was an abductive analysis (Timmermans and Tavory, 2012) in which we moved to and from between our empirical material, the theoretical literature on promissory rhetoric and the dream metaphor.

We decided to focus our analysis on the epistemic dimensions of promissory rhetoric, thus the promises related to new information and revolutionized knowledge production. We made this choice because many promissory rhetorics surrounding data-driven technologies are related to promises for new information (e.g., Andersson, 2008; Mayer-Schönberger and Cukier, 2014; Stevens et al., 2018). This meant that we did not look at sensemaking of promises related to cost saving or quality improvement as we assumed that many of these promises are implicitly related to epistemology. Hopes for more efficiency and cost-reduction of healthcare at least assume that the analytics are correct and produce relevant information.

During our analysis, we paid particular attention to instances where respondents shared personal experiences. These were the moments when a respondent became genuinely enthusiastic, included a touching personal story or changed their tone and began telling why something mattered or what they hoped for or dreamed about. We also looked for “nightmares” as sometimes it was easier for respondents to describe their fears than their hopes and wishes. In addition, because we had to deal with many interview transcripts, we built a simple text-mining tool to search for specific keywords, such as dream, wish, hope, personal, nightmare, story. This helped us to find key passages more easily.

We used the metaphor of dreams to analyze the material. We asked ourselves: who is dreaming this and why? What is the dream a response to? What are the related nightmares? What are the consequences of the dream? What does the dream lead to? This process eventually led to descriptions of four overarching data-driven dreams concerning healthcare information. The overarching dreams present four ways in which overarching promissory rhetoric is linked to the professionals’ longstanding frustrations, ideas, values and experiences in the healthcare sector. We presented the results to key respondents as a member check.



## Results

During the interviews, the respondents could suddenly switch tone and begin to “dream”. Discarding managerial-speak with its abstract statements on the importance of data and generic talk about “progress”, “improvement” and “revolution”, they began telling personal stories to explain why data-driven technologies were important to them. This is the story of a Spanish lawyer who specializes in data protection regulation.

*“For example, my daughter illness [...] They [medical doctors] told me that the one doctor who has seen the most cases like my daughter’s [...] had seen only four in his whole life! If you can’t gather information on all the people who have the same disease, you have nothing to work with! It’s so important that people know that we must share information because only then can you generate a pool of knowledge that allows you to compare one situation to many other similar situations. Big Data lets you do this” (Interview lawyer, Spain).*

Full of emotion, the lawyer explained that Big Data technologies could have helped cure his baby daughter who had a very rare disease. His daughter was lucky and eventually recovered, but the experience was horrific. He hoped that other parents and children did not have to go through the same. He saw innovative Big Data technologies as the solution and this meant that every day he worked to make safe data sharing possible.

Most respondents based their data-driven dreams on particular experiences with illness and health or longstanding frustration with the organization of healthcare. These frustrations often had to do with electronic patient records, the lack of data sharing and the difficulty of obtaining information. For instance, the respondents were frustrated that healthcare professionals in some hospitals still used “paper-based patient records” (interview research manager, Spain) and “fax machines” (interview professor in health informatics, Ireland). In these cases, the respondents linked the promise of data-driven technologies to their frustration, which became a motivator to revolutionize data use through new technologies.

To our surprise, we saw many similarities between the dreams of the professionals, even when the respondents regarded their dreams as individual, particular or “only” a personal anecdote. In the following, we describe four ways in which the professionals made sense of the data-driven promises using the dream metaphor. In each section, we elaborate on the motives and ways professionals link data-driven technologies to particular experiences and frustrations. We also highlight the values that are implicitly or explicitly prioritized and the considerations made. To set the tone, each section starts with our own characterization

of the relevant dreams and related nightmares, based on our own interpretation of the various stories we encountered in our interviews.

## Dreams of being seen

---

The dream: you are a patient entering a hospital. The counter staff greet you warmly. They already know that you are coming, remember you from a previous visit and point your way to the consultation room, even without asking! The doctor takes their time with you, listens properly, shares their thoughts and uses the best available evidence, tailored to your needs and displayed on a fancy computer screen that you can see too. You do not feel like any old patient. You feel that the doctor really sees you. Both of you look at the ways in which you can be treated and supported. At the end of the consultation, you feel great, you feel supported and are confident about the chosen treatment.

The nightmare: you are one of a crowd of patients in a hospital waiting room. The hospital feels like a big factory. Everyone is in a hurry and no professional makes eye contact. Once in a while, a doctor peeks out of a consultation room, and calls a name... another name... and yet another name. Eventually, they call yours and you follow the doctor into the office. The doctor asks a few general questions and quickly moves on to the usual procedure, based on the usual evidence. You get a standard folder explaining what you should do. You do not feel seen. You feel invisible.

---

In dreams of being seen, the respondents hope to see a data-driven future that tailors healthcare information to patients in their own context. Dreams of being seen were considered important when respondents talked about their own healthcare experiences or that of family and friends. As for the Spanish lawyer mentioned above, these dreams express the hope for diagnoses and treatments that are closely tied to individual patient characteristics. Sometimes the importance of tailoring patient care became more apparent through a professional's fears and worries. The "nightmares" depict the fear that they or their loved ones are not being seen and they will lose their identity or particularity. These nightmares portray a healthcare system that does not adapt information to the patient's situation, leading to "incorrect labeling of people" (interview chief medical information officer, the Netherlands) and other sorts of misjudgments.

Similar expressions for optimizing individual patient care have been around for decades but have not been broadly realized (e.g. Prainsack, 2018; Tutton, 2012). In the interviews, respondents often explicitly related the promises of data to their wishes for personalized care. For example a computer scientist from Sweden said, "I think it's the basis of it all, at least here in Sweden, of lots of expectations: personalized medicine, tailored medicine, optimized healthcare based on individual needs, and so on" (interview computer scientist, Sweden). Ideas initially popular in genetics and genomic research seem to have moved into the realm of data. Genetics research promised individualized treatments based on our DNA. The hope now is to personalize care by including all sorts of lifestyle and environmental data (Prainsack, 2018).

Dreams of being seen inhabit a widely shared sense of dissatisfaction with the general level of information used in the healthcare sector. Medical professionals often work with general guidelines or make decisions based on research conducted on specific patient populations (often “White” male, without other medical conditions) (e.g. Epstein 2009). Care information is thus about certain groups of patients, averages and usual responses and this restricts its application. Patients often receive a particular treatment because it worked for many other patients or because it is the protocol to do it that way. This means that healthcare professionals are often unable to predict exactly which treatment will work best for which patient.

These dreams prioritize values such as individuality, personalization and uniqueness. The assumption is that everyone is different, particular and out of the ordinary and should be treated as such. At the same time, the dreams tend to overlook such values as collectiveness, generalization and universality. However, knowledge and insights can be obtained because a particular situation is compared to a general situation. Knowledge can be obtained because something has been seen before. Seldom about individuals, it may have more of a focus on groups. The data-driven techniques promise individualization but in reality, deal more with specific groups. The underlying hope is to make more precise comparisons based on data and analytics.

An important consideration of this dream is that individualized information is more important than data protection. Take, for example, a health insurer from France. During his interview, this Frenchman began dreaming about a pre-diabetes test: he linked ideas about data use for individualization to his own case of being, in his words, “slightly overweight” and thus at risk for diabetes.

*“I’m not sure but I don’t think I’ve ever been tested for pre-diabetes. I’m in my early fifties, slightly overweight. So, I could be at risk of developing diabetes. Look at me, I don’t think I’m obese (but) truth is, nobody really cares. I’d love to have somebody tell me that I need to take a pre-diabetes test [...]. I’d love to have somebody tell me if it’s right for me to do that. And of course, you’d need a set of data, otherwise, it’d just be an opinion. [...] Maybe [...] checking if I’m at risk shouldn’t just look at scientific data but at my behavioral data as well. Most of my spending is by credit card. I don’t really buy fast food or stuff like that very often. But if I did, maybe using [behavioral] data would let somebody catch it and get back to me. So, I’ve been a bit specific, but that’s what I’d like to see in the future” (Interview health insurer, France).*

He says that currently “nobody really cares” about his health and lifestyle. He would love to have a healthcare professional bring together his data, physical and behavioral characteristics to determine whether he needs specialist treatment. He does not mind disclosing personal information about his shopping patterns (credit card behavior), because he thinks it is valuable to have information about his specific risk for developing a disease.

## Dreams of being in control

---

The dream: you are a doctor seeing a patient in the outpatient clinic. The patient explains his problem. You decide to open the medical dashboard on your computer. You see a comprehensive overview: five suggestions for the most likely diagnoses, but also highlights of the patient's history. You decide to click on the top diagnosis. Now, you see how the system came to this particular diagnosis and you get a good overview of treatment options and chances that they will cure the patient's problem. These overviews make you feel in control. You have all the necessary information at hand and can use it to make treatment decisions.

The nightmare: you are a medical doctor trying to treat a patient. You listen to their story and try to make treatment decisions. But you cannot get a grip on the situation: new side effects constantly appear, and new symptoms develop as soon as you have cleared a symptom somewhere else. You cannot manage the disease and the patient's condition deteriorates. You look up the scientific evidence, medical guidelines and ask for advice from other professionals. You try to control things, but nothing works. Everything keeps on breaking down around you and you do not understand why.

---

In dreams of being in control, respondents hope for a data-driven future in which healthcare information is known and controlled and decision-making can be optimized. They dreamed, for example, about developing medical dashboards that “doctors can use and draw conclusions from” (interview patient representative, the Netherlands) or about other useful ways to present information. These dreams were also apparent in comments on using specific data, methods and analyses. For instance, a medical researcher from Ireland stresses the importance of data quality when she argues for using proper strategies to obtain good quality data:

*“Coming from the health research world, I would see Big Data as [...] tapping into large cohorts of populations and patient studies [...] to find data in the proper quality with the proper integrity in place [...] and then re-use them” (Interview medical researcher, Ireland).*

Sometimes the respondents' worries expressed the importance of being in charge of data and information. In such nightmares, professionals fear losing control over healthcare decisions. For example, because healthcare decision-making is outsourced to self-learning systems and knowledge production is “black-boxed and so locked up too much” (interview professor in public health, Sweden) for human intervention.

Control dreams inhabit a widely shared sense of dissatisfaction with the uncontrollable and out-of-reach character of information in the healthcare sector. Diagnostic and treatment processes are complex collaborative activities, with medical doctors collecting information from various sources (Balogh et al., 2015; Van Baalen, 2019). Doctors rely on patients, their family members, technological experts, laboratory staff, and their colleagues. They also rely on medical guidelines, scientific studies and all sorts of technologies to develop a diagnosis or treatment (Van Baalen, 2019). Through an iterative process, medical professionals try to gather information and reduce uncertainty as much as possible. However, diagnostic and treatment processes are not easy for medical professionals under time constraints. The gathering process is complex and sometimes little evidence is available (Balogh et al., 2015; Grote and Berens, 2019).

Data-driven technologies might enhance decision-making (diagnostic and treatment) processes by offering new kinds of certainty (Grote and Berens, 2019). For example, a Spanish lawyer explained that he hoped that data-driven technologies would enable new sorts of information that let us “really understand what is happening”:

*“I believe that we have the technical means that allow us to gather the biggest packets of information that have been blurry so far. I mean, they were hard to collect, put together and analyze consistently. Now, we can get information from very different sources. We can run algorithms that let you interpret what’s happening. This was so hard to do before, when we didn’t have all those, you know, computers and algorithms. [Now we] really understand what’s happening” (Interview lawyer, Spain).*

Similar desires for control of health and illness have always been present in healthcare but now they are fueled by the emergence of data science as a field (Ribes et al., 2019; Stevens et al., 2020) and repeat many scientific promises: that with certain methods and technologies we can uncover the truth about illness and disease, and we will be able to intervene successfully.

These dreams prioritize such values as consistency and control. These promises often seem to portray a one-sided view of data, evidence and information. What seems to be lacking is an acknowledgment that information cannot always be added together coherently; instead it can lead to opposing insights and suggest multiple interpretations (cf. Mol, 2002). These dreams neglect the values of inconsistency and flexibility. The idea is that all the information in the world will provide a complete, coherent overview, whereas all this gathered information might also produce a cacophony of insights.

The respondents consider human authority and control highly important in these dreams. Some shared examples of self-learning systems could produce information that could conflict with the professional assessment of a medical doctor. In such a case, humans would have final authority as the respondents felt it was important to be in control, to understand where information comes from and how it is obtained. A Dutch think tank employee explained:

*"I don't believe in that doomsday scenario: that a self-learning system will take over from us and go its own way. Ultimately, as humans, we'll still be the ones who program it, who'll teach the device how it should learn" (Interview think tank employee, the Netherlands).*

## Dreams of timeliness

---

The dream: anyone can log into a [health risk warning] system to find out what disease(s) they might get in future. You are happy with this advance because it gives you time to take all sorts of precautions. You can change your lifestyle and watch out for early signs of disease. It means that you can catch it early so that it will have less of an impact. Because early intervention is possible, it could stop some diseases from happening at all.

The nightmare: you have no idea what will happen to your health in future. So many diseases lie in wait, and you do not know if and when a particular disease will hit you. You try to watch for the signs and take precautions, but what will happen and what direction to take is a mystery. This makes you insecure and hesitant to do anything.

---

In dreams of timeliness, the respondents talked about real-time analytics that would remove any delay between data collection and analysis. They shared their hopes for improved prediction and prevention, so that professionals and patients could know what will happen in the future and act accordingly. A policymaker from Sweden explains:

*"It's a way of make a healthcare dream come true: not only treating patients when they have fallen ill, but preventing them from even getting ill" (Interview policymaker, Sweden).*

Again, the professionals' fears and worries expressed the importance of timeliness. In these nightmares, the respondents are taken surprise by unexpected disease or are too late to intervene.

These dreams inhabit a widely shared sense of dissatisfaction with the lack of timely information in healthcare. Indirectly, this dream critiques current scientific methodologies, such as healthcare experiments, as they often take a long time and provide information that is

not always easily applicable to everyday healthcare (Nicolini et al., 2008; Stevens et al., 2020; Timmermans and Berg, 2003).

During the interviews, the respondents explored the opportunities to act on timely information. For instance, a professor in health informatics from Ireland thought about how overweight people could be nudged into adopting a healthier lifestyle:

*"We can make interventions earlier; we can try and identify people's risk of conditions quicker. [...] If we know someone is overweight, we can monitor what they are buying. If we see that they are buying really poor food, if we know that they have a very sedentary lifestyle [...] then we could actually intervene and try and encourage them to change their lifestyle a small bit. Try and hopefully prevent them from getting diabetes and thus, a lifetime of difficulty. It relates to education [...], to policies [...]. But the potential is there! [...]. Why could we not develop wearables and take that data and start to really look at it? Why could we not allow those patients to manage themselves better at home? We have machine learning systems that are keeping an eye on them and flag up when they are potentially going off track and (when) we need to intervene. Then the nurse or the doctor, the general practitioner, can pick up the phone and [...] check up and make sure you're ok. I think there's tons of potential like that. How far away are we? Five, ten, fifteen 15 years but, definitely, you can see where it's going" (Interview, professor in health informatics, Ireland).*

The quote illustrates a central idea behind the hope for timely information: quick identification of a person's health risks facilitates intervention. Here the underlying assumption seems to be that past knowledge, grounded in enough data, could let us predict the future with near certainty (Busch, 2016).

These dreams prioritize such values as foresight and timeliness. It is important to know that something will happen, but at the same time, values such as unexpectedness, unawareness are neglected. The dreams assume chronological trajectories and that intervention is possible when things are known sooner. However, as others have argued, the complexity of our social world makes it hard to fix the past and predict the future (Busch, 2016).

Although these dreams consider both timeliness and the opportunity to act important, they neglect the fact that sometimes nothing can be done, given a particular prediction. This means that people are given the obligations to act and to know, which leads to all sorts of new responsibilities (Mol, 2006). People can be turned into "partial patients" or "patients in waiting" (Harris et al., 2006) when no action plan may be available.

## Dreams of connectedness

---

The dream: you are a doctor facing a patient with a rare disease. You do not know what to do because you have never seen a case like this before. Luckily, you can use data-driven technology. You can look up the patient's complete medical history and download data from the patient's wearables. You can compare this information on a database of patients with similar diseases worldwide. This connected information enables you to diagnose the patient and decide on treatment.

The nightmare: you have the same patient sitting in front of you. You do not know what to do but cannot find good information. You find a study here, something interesting over there, but cannot connect these together. This means that you must tell the patient that you do not know how to help him. You are at a loss.

---

In dreams of connectedness, the respondents frequently talked about connecting data and datasets and stressed the importance of data sharing. They emphasized the distributed nature of data in the current healthcare sector, where many organizations collect, store and use only their own data. A computer scientist from France explained the desire for more general sharing:

*"I would like generalized medical records, so I know when I go to the hospital that the doctor in front of me has immediate access to all my data and can treat me"*  
(Interview data scientist, France).

The importance of being on time sometimes became apparent in the respondents' expressed fears and worries. One clear example is the quote by the Spanish lawyer with his critically ill daughter. He nearly lost her due to the lack of information sharing in this case of a very rare disease.

These dreams confront the fragmented, distributed nature of knowledge in the healthcare sector, where data are stored in every imaginable format at numerous locations under the control of many different people (Nicolini et al., 2008). Many respondents shared their frustration with the lack of digitization of all sorts of healthcare information. A medical researcher from Ireland dreamed about a national system that would enable easy sharing of electronic health records:

*"We don't have electronic health records. At the individual disease stage some hospitals have electronic databases of patients, kind of patient case-report forming. But they are not part of a national link, [they're not] joined-up. Type in your number and all your health records are in the [system]. That's the stuff of dreams"*  
(Interview medical researcher, Ireland).

Some respondents take this idea even further and see the potential of connecting medical data from hospitals to patient wearables and other data-gathering technologies that gather



even larger datasets. Here the underlying assumption seems to be that bringing together different data sources creates greater mass, leading to improved information overall. This is reminiscent of a “modernist discourse” that assumes that data are “out there” and only need be found and brought together to come to new information (Stevens et al., 2018).

These dreams prioritize such values as harmony and sharing but neglect such values as incompatibility and competition. While it might not be straightforward as even within hospitals, many groups have diverse interests that do not necessarily align (Tamuz et al., 2011).

The respondents considered many aspects of interoperability and standardization important in these dreams. They mentioned technical, semantic and organizational functionality, as well as international standardization of regulations and procedures and norms and values that need bringing together. However, these efforts at connection disregard the particularity and context in which data are gathered and analyzed (boyd and Crawford, 2012).

## Discussion

*“The critic is not the one who debunks, but the one who assembles. The critic is not the one who lifts the rug from under the feet of the naïve believers, but the one who offers the participants arenas in which to gather” (Latour, 2004: 246).*

In this chapter, we set out to explore the personal sensemaking of diverse professionals working on data-driven initiatives in healthcare. The aspect of personal sensemaking is largely absent from both techno-optimistic analyses of the promises data-driven technologies bring to the health domain and many critical analyses of the rhetoric, buzzwords and expectations that surround such technologies. Applying the metaphor of dreams, we explored data-driven dreams about healthcare information in order to highlight the motivations (why are people mobilized by the promises?), values (why do they find the promises important?) and considerations (what are the trade-offs that they make?) in pursuing data-driven healthcare care. We highlighted four dreams in particular: dreams of being seen, of being in control, of timeliness and of connectedness. Here we reflect on the metaphor of dreams and consider how bestowing precise analytical attention to personal sensemaking processes can contribute positively to the CDS literature.

The metaphor of dreams helped us understand how professionals working on data-driven initiatives make sense of the grand promissory rhetoric surrounding data-driven technologies. We saw our respondents bring grand promises back to the sphere of the personal

that they can influence and understand. The respondents link the grand promises to their own experiences with health and healthcare, which could include general, scattered, slow or uncontrollable information. They also connect the grand promises to desires and needs that are arguably fundamentally human: being seen, feeling connected, feeling secure, and feeling protected against harm. Our analysis of personal sensemaking through the metaphor of dreams has produced new insights into the affective, relational and attentive dimensions of data-driven promises.

CDS scholars have done valuable work on exposing the harms and assessing the many consequences of data. CDS scholars open up data processes by showing, for example, how data is shaped, what work is required to obtain data (analyses) and how they produce specific knowledge (e.g. Pine and Bossen, 2020; Prainsack, 2019; Shelton, 2020; Whitman, 2020; Ziewitz, 2017). They also reveal the patterns of power, for instance, by exposing forms of data governance, control and exploitation (e.g. Micheli et al., 2020; Tanninen, 2020). And lastly, CDS scholars pay meticulous attention to a wide variety of negative effects and how these (re)produce forms of discrimination and inequality (e.g. boyd and Crawford, 2012; Lee and Larsen, 2019; Tubaro et al., 2020; Cruz, 2020).

While these studies produce a range of valuable insights, they also feature an uneasy emphasis on the negative effects of data-driven technologies. CDS is excellent at critically assessing technological developments and has created a sharp set of analytical sensitivities for probing the range of ethical, epistemological, political and ontological complexities and tensions that surround data-driven technologies. However, as in any discipline, there are also blind spots, as is exemplified by the personal sensemaking of such actors as legal, governance and ethical experts, healthcare staff and data scientists. This means potentially neglecting the complex practices surrounding data-driven promises in healthcare that are already full of affectivities and normativity. It also means potentially introducing overarching normative elements from the outside that do not do justice to the normativity in the studied practices themselves (Zuiderent-Jerak, 2007).

A consequence of this criticism from the sidelines is that many research insights do not actually reach the professionals working in the field on data-driven initiatives. We have argued previously in this journal (Stevens et al., 2018) that insights from CDS do not seem to infiltrate the healthcare domain as there is little attention for the constructed nature of data and its epistemological assumptions. Discussions about the permissibility of data-driven techniques are for the large part outsourced to ethics and limited to such themes as privacy (Wehrens et al., 2018). Moats and Seaver (2019) describe similar “divides” between CDS and data science practice, highlighting that computer scientists often raise concerns about their disciplines but tend to base these on their own normative

assumptions “about what constitutes ‘robustness’ or even ‘research’ in the first place” (Moats and Seaver, 2019: 2) thereby missing other CDS concerns.

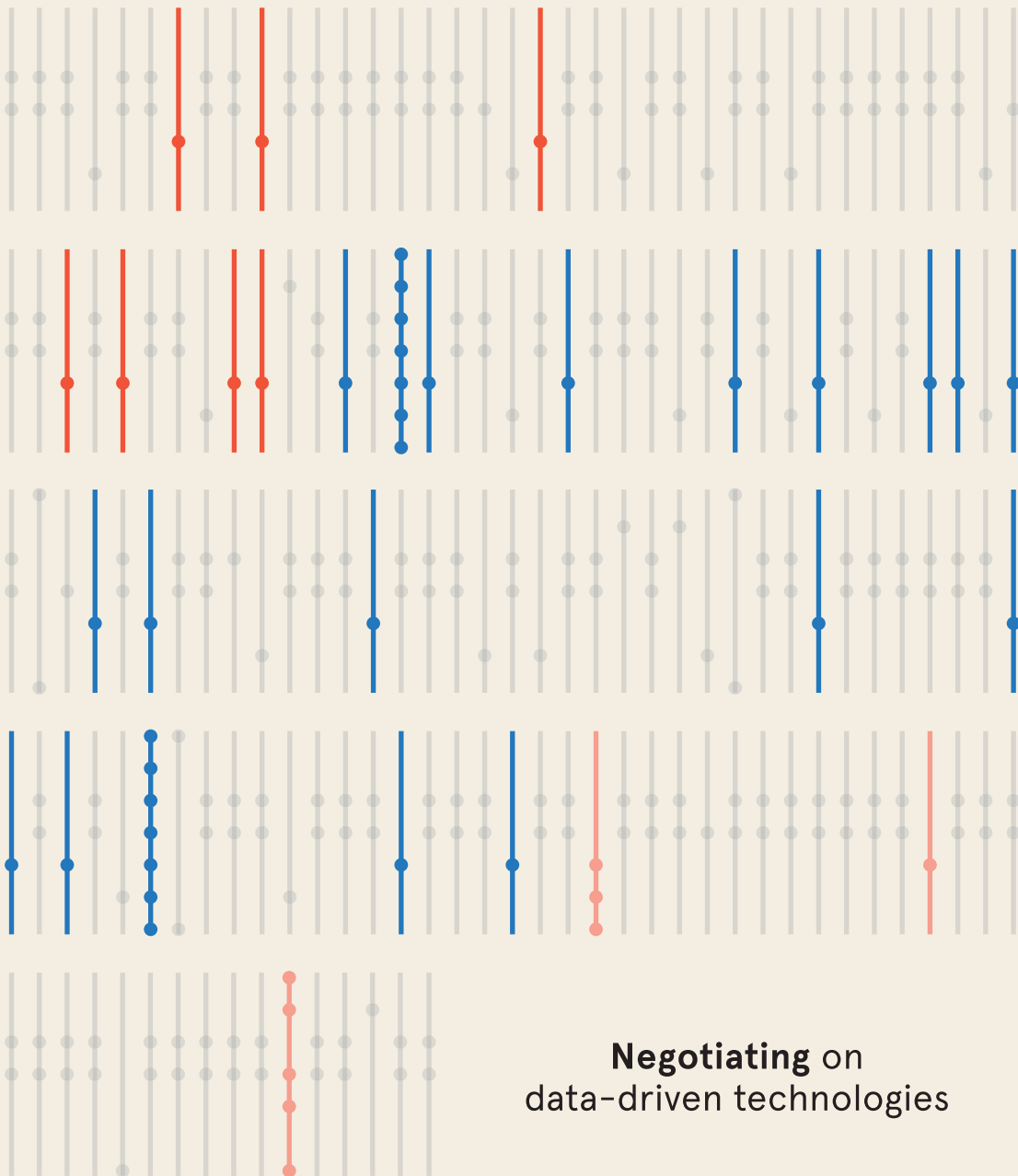
The divides between CDS and other fields are not natural, nor are they inevitable. On the contrary, they are shaped and can be negotiated on in concrete practice (Galison, 1997; Stevens et al., 2020). There is an urgent need for constructive approaches and creative concepts to overcome the differences, to break people away from abstract theories and (statistical) methods that can obscure opportunities for collaboration (Stevens et al., 2020) and can stimulate communication between fields without rehearsing the distinctions between actors, practices and methodologies (Moats and Seaver, 2019). However, this is not easy to achieve, as our initiatives could very well reinforce the divides and differences, if only inadvertently (Moats and Seaver, 2019).

In conclusion, if the above interpretation is recognized, it leads to an important question: how to proceed? Given the need for increased learning and understanding “between” our insights and those in healthcare practice in order to “create hybrids” (Moats and Seaver, 2019), we argue that it is vital to study data-driven practices and shed light on the affective engagements through which data-driven technologies and initiatives are formed. Our way of studying these dimensions has been through the metaphor of dreams, but there are many more options. Most important is that we dive into the complexities, interdependencies and tensions involved with data in practice to help care (Puig de la Bellacasa, 2017) for “responsible” data-driven technologies.

Published as:

Stevens M, Wehrens R and de Bont A (2020) Epistemic virtues and data-driven dreams: on similarities and differences between data-driven and medical practices. *Social Science & Medicine* 258(113116): 1–8.

# Chapter 5





## Introduction

Data-driven initiatives aim to analyze large volumes of data from varied sources to improve healthcare delivery by predicting future health risks and treatment responses (Mayer-Schönberger and Cukier, 2014; Murdoch and Detsky, 2013). Examples of such initiatives are machine learning applications that analyze medical images (Sample, 2018) or predict readmissions on intensive care units (PacMed, 2019) and artificial intelligence systems that assist in the diagnosis of cancer (Somsashekhar et al., 2017).

Such initiatives often come surrounded by promissory discourses that emphasize how large-scale or innovative data analysis will result in valid information to be used to enhance healthcare provision. These discourses result in a strong positive rhetoric about data-driven initiatives and as such drives their implementation. However, more cautionary discourses frame these initiatives in more critical terms and see the new initiatives as only a possible (or even a risk) to current epistemic practices in the healthcare sector (Stevens et al., 2018).

As fields originating in different disciplines, medical and data scientist communities have diverse epistemic practices, which come together in data-driven initiatives. Epistemic practices guide how members of a field propose, communicate, evaluate and legitimize knowledge. These epistemic practices are part of particular epistemic cultures (Knorr Cetina, 1981) that can be described as sets of specific norms, values, beliefs and traditions, that are “bonded through affinity, necessity and historical incidence” (Knorr Cetina, 1999: 1). This means that epistemic cultures are known for specific activities for reasoning and establishing evidence, thereby determining what and how we know in communities.

The literature on data-driven techniques seems to either downplay or overemphasize the epistemic differences between data science and medical fields. On the one hand, data science scholars and some medical professionals easily dismiss their differences (e.g. Amato et al., 2013; Mayer-Schönberger and Cukier, 2014; Murdoch and Detsky, 2013). They imply that it is only a matter of time before the methods and standards of data science become part of established medical practice, thus suggesting an ultimate trajectory of convergence that downplays distinctions between their epistemic practices.

On the other hand, scholars are critical of the evidentiary claims of data scientists and primarily stress the epistemic differences. Some medical professionals explain in detail where and how the methods of both communities differ by theorizing on the specific role of data-driven approaches. They contrast grand concepts like causality and correlations and argue that data science can, for example, only be used to generate hypotheses and to explore valuable research directions (e.g. Khoury and Ioannidis, 2014). They argue for the

specific use of data-driven approaches by highlighting the distinctions in terms of methods and evidentiary standards.

The aim of this chapter is to study the concrete negotiation of differences in the epistemic cultures of data science and medicine without assuming either absolute incommensurability or deterministic complacency. Earlier studies about technological innovations have shown us that comparing (downplaying or emphasizing) the distinctions with the status quo is not helpful (Janssen, 2016; Smits, 2002; Van Lente, 2012) and we argue that it is more interesting to study where and how the epistemic cultures of data science and medicine overlap and differ. Such an approach is more likely to capture all the nuances, efforts and workarounds in practice. In addition, a more detailed empirical understanding of the negotiations in practice can also provide useful insights into how practitioners from diverse fields find ways to work with each other despite epistemic differences. This has broader implications for interdisciplinary knowledge practices that are increasingly more common and more expected (e.g. large European projects) (Rathenau, 2018).

We conducted a case study of a data-driven initiative in the context of psychiatry. Psychiatry is a particularly interesting case to study the introduction of data-driven initiatives as this field is characterized by considerable uncertainty relating to disease ontology and treatment effects. Moreover, it relies primarily on narratives of patients and qualitative questionnaires to make sense of patients' conditions and guide treatment decisions. This means that uncertainty related to data collection and use, which is also part of other fields of medicine, is amplified in the field of psychiatry.

The chapter is based on an ethnographic, empirical study of the development and implementation of a hospital-based data-driven initiative in the Netherlands. Within this initiative, data scientists were brought into a psychiatric hospital department to develop prediction models of patient outcomes based on machine learning techniques. We approach this initiative through the sensitizing concepts of "epistemic virtues" (Daston and Galison, 2007) and "trading zone" (Galison, 1997). The notion of epistemic virtues enables us to ethnographically focus on the differences in epistemic cultures, while the concept of trading zones allows us to zoom in on how the differences are negotiated. This leads to the following research question: *how are epistemic differences negotiated by data scientists and psychiatrists in a hospital-based data-driven initiative?*

In the next section, we sketch the theoretical background and elaborate on epistemic differences by introducing the concepts of epistemic virtues and trading zones. We explain how this combination helps analyze the practices in the initiative. This is followed by the case description and methodology. In the results, we introduce the data-driven initiative



and present two cases from our fieldwork that illustrate the process of trading epistemic differences. The discussion concentrates on the role of epistemic virtues and how they play a role in interdisciplinary cooperation.

## Theoretical framework

### Epistemic differences

A central theme within STS has been critiques and reflections on the idea of “epistemic unity of the sciences” (Galison and Stump, 1995). In this tradition, Knorr Cetina (1981;1999) developed her work on “epistemic cultures”. She is known for her ethnographic comparison between experimental high-energy physics and molecular biology. She shows that knowledge is created in diverse scientific cultures and this results in a wide array of scientific practices and preferences that all coexist. Building on this body of work, we conceptualize data science and psychiatry as two different, but not necessarily incommensurable cultures of knowledge with their own epistemic practices.

This conceptualization of data science and psychiatry as diverse epistemic cultures is grounded in literature. Scholars investigating data science have, for instance, pointed out how the field is characterized by an epistemic tradition with a specific history that is different from the epistemic culture described in literature about psychiatry. It is known that data science is a relatively new field (Baru, 2019) that grew with the capacities to gather and analyze large amounts of data and the increasing “datafication” of aspects of our life (Mayer-Schönberger and Cukier, 2014). It originated from computer science, statistics and mathematics, but more explicitly engages with “real” data under “real-world” constraints (Baru, 2019). A strong positive rhetoric and commercial successes led to the increased realization that real data are a valuable commodity and new methods, infrastructures, technologies and skills were being developed to handle these data (Leonelli, 2014). The field is described as relatively a-theoretical, as it is a general approach that can be applied to analyze data drawn from a wide variety of fields and domains (Ribes, 2019).

Epistemically, the field of data science assumes that there is a (be it, complex and multi-faceted) real world or reality that can be better or more completely captured by real-world data (Stevens et al., 2018; Leonelli, 2014; Ribes, 2019). Thus, the field assumes that there is a strong connection between the real-world and data. This sentiment leads to the suggestion that we can know more and act better from the careful grouping and analysis of real-world data and extrapolating them into the future. This commitment to these data means that this field is (in principle) respectful of any statistical relationship between two

data values and accepts that some relations might never be understandable or explainable (Leonelli, 2014).

The field of psychiatry has a longer history than the field of data science. Since the twentieth century, the field of psychiatry has undergone numerous and by no means monolithic transformations (Rüppel and Voigt, 2019). During this time, psychiatry has been influenced by the evidence-based medicine (EBM) movement that resulted in a preoccupation with randomized controlled trials (RCTs) and experimentation (Timmermans and Berg, 2003; Rüppel and Voigt, 2019) and also by the “molecularization of the “medical gaze” (Rose, 2007) that led to studies into genetics and biomarkers that aimed to localize diseases within the body (Rüppel and Voigt, 2019). However, the EBM-movement had more impact in other fields, as measurements could be obtained under more controlled conditions and with more exactitude and accuracy. Therefore, psychiatry can still be characterized by considerable uncertainty relating to disease ontology, treatment effects and strong reliance on patient narratives.

Epistemically, the field of psychiatry assumes that diverse genetical, neurobiological, environmental, biographical factors and the interactions between them lead to the development of psychiatric diseases in some people (Rüppel and Voigt, 2019). Thus, the field assumes that there are (strong) causal connections, but they are complicated and often unknown. This sentiment results in diverse kinds of research that aim to capture relationships between events and outcomes. These studies and their evidence are consequently ranked and evaluated based on their quality and conventions (Vidal and Ortega, 2017). The preference for specific research methods means that the field is critical about scientific methods (e.g. large observational studies and RCTs are preferred above individual case studies). This contributes to gaps in evidence that need to be bridged by clinical interpretation in practice (Rüppel and Voigt, 2019).

This literature suggests that both data science and psychiatry assume a strong and complex connection between events in the world and data that are gathered about them. These data can be analyzed and are, in principle, a sound basis for decision-making. However, there are differences in the sort of data and methods that are used and how they are valued.

### **Epistemic virtues**

In order to understand where and how the epistemic cultures of data science and psychiatry differ and to study how these differences are concretely negotiated in data-driven initiatives, we build on the concept of epistemic virtues. Research in philosophy of science

and STS have shown that epistemic norms and values play a role in scientific communities and guide what is perceived to be normal and acceptable (Kuhn, 1962; Latour, 1987).

Epistemic virtues can be conceptualized as epistemic norms and values that are internalized and acted upon by data scientists and psychiatrists (Daston and Galison, 2007: 40–41) and are one of the ways of studying epistemic norms and values. Both data science and psychiatry have internalized norms about, for instance, certainty, representativeness and objectivity. These can be understood as virtues that are acted upon in specific ways, for example, in the judgements that medical practitioners pass on one another's work. They do not use the word "virtue" or "vice", but their praise or blame often relates to qualities of work that they consider "good" or "bad" (Paul, 2011: 7). The concept of epistemic virtues thus helps us to look ethnographically at precisely those moments when specific judgements are made by data scientists and psychiatry practitioners. Epistemic virtues enable us to interpret such moments as not merely a methodological discussion, but as examples that signal underlying differences in epistemic cultures.

In contrast to approaches that tend to list epistemic virtues considered relevant for knowledge gathering (e.g. Marcum, 2017; Pigliucci, 2017), our use of epistemic virtues stresses the dynamic and social character of epistemic virtues (and hence also epistemic cultures). This understanding of the notion emphasizes that: (1) epistemic virtues are actively constructed and continuously re-evaluated and, (2) epistemic virtues are situated, which means that people, often together, always give meaning to epistemic virtues in specific circumstances.

We are inspired by the work of Daston and Galison (2007) around the notion of "objectivity". They showed how the meaning of the epistemic virtue "objectivity" was (re) constructed throughout the years and showed how objectivity was understood differently by scientists in the eighteenth, nineteenth and twentieth centuries. Their approach inspired us to look at the multiple interpretations that are given to epistemic virtues and what this says about sameness and differences of two epistemic cultures. As a useful addition to the work of Daston and Galison (2007), we suggest that the notion of epistemic virtues is not only useful to study longitudinal changes in scientific development, but can also be fruitfully applied to study the in situ negotiation of differences in epistemic cultures (Knorr Cetina, 1981).

## Trading zones

While the notion of epistemic virtues enables us to ethnographically focus on the differences in epistemic cultures through concrete moments of judgment, the concept of trading zones allows us to zoom in on how these differences are negotiated. We frame

the data-driven initiative as a trading zone (Galison, 1997). Galison (1997) also challenged the “epistemic unity of the sciences” in his work about trading zones as he was intrigued by the extraordinary variety and disunity of scientific languages and practices (Galison and Stump, 1995). He analyzed how distinct communities in physics – such as theorists, experimentalists and engineers - create in-between vocabularies that facilitate communication and alignment of activities (Galison, 1997). These intermediating languages can range from simple (interlanguages) to complex (“pidgin”) and eventually a shared language can emerge (“creole”) (Collins et al., 2007; Galison, 1997). These languages make it possible to interact and exchange goods despite differences and without homogenizing the inherent diversity in their communities (Galison, 1997).

Galison’s notion of trading zone is used as a tool to analyze a wide variety of interactions between different communities, ranging from cross-boundary interactions between surgeons and engineers (Baird and Cohen, 1999), ways in which team members with different backgrounds cooperate in a marketing firm (Kellogg et al., 2006) to communication between NASA engineers and their subcontractors (Vaughan, 1999). This body of (ethnographic) work continuously stresses that the different communities do not meet each other “with gaping incomprehension” (Harmen and Galison, 2008: 568), but that members of the diverse communities coordinate their actions temporarily and locally. They navigate their differences in language and culture only as needed (Kellogg et al., 2006; Galison, 1997) and this is exactly what gives science its strength and coherence and underlies the experience of scientific continuity (Galison and Stump, 1995).

The trading zone literature pays attention to the linguistic and material components that help cross-boundary interactions. Galison emphasized the importance of language right from the start. Later, Collins et al. (2007) expanded on the linguistic understanding of trading zones with their work on “interactional expertise”. This concept highlights two things. First, that not only communities but also independent third parties can gain interactional expertise in talking to both communities in some approximation of their language (Collins et al., 2007; Collins and Evans, 2008). Second and more importantly, interactional expertise emphasizes the social processes and tacit knowledge that are passed on through language and that are important to facilitate exchange (e.g. Epstein, 1996). This work highlights that socialization is important for learning another culture because through day-to-day immersion, people learn and understand the rules that cannot be written down (Collins and Evans, 2008; Wehrens, 2015).

Besides the linguistic components, there has been a focus on material, “boundary” (Star and Griesemer, 1989) objects that facilitate exchanges largely in the absence of linguistic interactions. The objects facilitate trading because they are “plastic enough to adapt to

local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites" (Star and Griesemer, 1989: 411). The objects can mean something distinctively different to both parties and help interactions because they do not vitiate the communities' separate projects (Collins et al., 2007).

In this chapter, we focus on the negotiation of epistemic differences and while language and objects can help to facilitate exchange, much is still unknown about the role of epistemic virtues in trading zones. Galison (1997) remarks that communities can have diverse norms, values and virtues. According to Galison (1997: 401, 807), the communities persistently try to incorporate the virtues of the other. However, a more in-depth exploration of the role of epistemic virtues in trading zones still has to be developed. We aim to contribute to this development by paying attention to precisely those moments in the initiative when specific judgments are made by data scientists and psychiatry practitioners. We argue that the focus on epistemic virtues as concrete moments of judgment can be a useful additional dimension in the empirical study of trading zones, next to linguistic and material dimensions developed by other authors.

## Methods

### Case description

This study focuses on a data-driven initiative by a psychiatric department in one of the largest Dutch University Hospitals. Combining outpatient and inpatient treatment, this department contains four specialist units that specialize in affective, psychotic and developmental disorders, acute and long-term care. The department treats approximately 2000 unique patients annually. There is a strong focus on research in the department as it is linked to various research groups, and many psychiatric practitioners are also involved in research. Studies focus, for example, on genetic predispositions, brain morphology, and risk factors that play a role in the development of psychiatric disorders. The department is well-established as a research center, both in the Netherlands and worldwide.

The data-driven initiative was initiated by the medical head of the department in 2015 as she set out to experiment with data-driven techniques to improve treatment for psychiatric patients, in addition to research initiatives already being conducted in the hospital. The department attracted data scientists who had worked for the major companies or were involved in data science start-ups. The initiative secured the help of an IT company and acquired funding from the Dutch government and other organizations. With these funds, the data scientists could start pilots, build structures that automatically supplied data and start working toward making care improvements for patients.

When we followed the initiative, the core of the data-driven team comprised four data scientists, that were supported by multiple students, six external data scientists and one data engineer. Students (e.g., medical, psychology) helped the data scientists with small projects. The six external data scientists were hired to work on short-term subprojects with specific goals in mind (they followed the funding). The data engineer worked for the university that was linked to the hospital and provided technical support.

At the time, the team worked on several data-driven subprojects. Besides presenting straightforward business analytics (e.g. how often is a certain medicine prescribed?), the team worked on developing decision-support systems that would aid psychiatrists by making predictions based on machine learning algorithms and experimented with the implementation of the machine learning models. Two major subprojects tried to predict which antipsychotic or antidepressant would be the most effective while producing the least side effects for certain patients. These projects were relevant because they would (1) avoid the current trial and error method, where psychiatrists must try several drugs before finding the right one; (2) reduce hospital stay; (3) be expandable to other psychiatric medication in the future.

## Data collection

Our empirical analysis builds on a combination of qualitative methods. First, MS observed the data-driven team between September 2017 and February 2018 (approximately 200 hours). After an initial meeting with the medical head of the department and the leader of the data-driven team in which we explained our research focus, we could meet and shadow the data-driven team. MS shadowed the data scientists while they were doing their analyses and followed them around to meetings in the hospital, taking field notes to capture the setting and interactions between data scientists, medical staff and researchers (Oldenhof, 2015). MS introduced herself and the research to all involved and made sure ongoing informed consent was obtained. Informal conversations helped build rapport and provided insights into the pertinent issues and tensions. The field notes were expanded as soon as possible and discussed with RW and AdB.

Second, MS conducted seventeen semi-structured interviews with nineteen people (directly and indirectly involved with the data-driven initiative) to explore particular topics in more depth: nine data scientists, three psychiatrists, two nurses, one data engineer, one data manager, one consultant, one hospital manager and one medical researcher. The interviews posed open questions about the initiative and the collaboration, such as: how do you go on collaborate with others? All interviews lasted 40–60 min and were recorded and transcribed ad verbatim. We asked permission for the interviews, the use of quotes and anonymized the material.

Finally, MS conducted an analysis of online and offline documentation about the initiative, such as presentations and newsletters. This resulted in more than 400 pages related to the initiative. An ethical waiver was obtained for the study.

## Data analysis

We began our analysis with open coding (Mortelmans, 2007). We performed general readings and highlighted notable passages in the data. For example, we marked passages in which practitioners from one community expressed that they did not understand the other community. MS conducted most of the analysis, but researcher triangulation ensured that key themes emerging from the analysis were discussed and refined (Mortelmans, 2007). Through several iterations of the analytical process, we focused on three key categories: (1) epistemic claims and ideas, (2) the negotiations and discussions between data-driven and psychiatry practitioners, and (3) the coordination of the meanings and ideas about knowledge production.

Although analysis began with open coding, the total process employed an abductive approach involving an iterative to-and-from between analytical themes and relevant theoretical concepts (Timmermans and Tavory, 2012). After identifying the literature on trading zones and epistemic virtues as crucially relevant for the analysis of the empirical material, we developed the analysis further through posing such questions; (1) what is traded here, by whom and where? (2) Which epistemic virtues are in play? (3) What does this say about the respective epistemic cultures?

In the end, our analysis focused on eighteen key practitioners involved in the initiative. Table 1 provides an overview of the practitioners per community, their role and data collected about them. The remaining people under observation had a minor role in the trading process. Their interviews helped to gain more information about the structure of the trading zone and details of the various projects. Due to time constraints and high workloads, not all psychiatric practitioners could be interviewed.

We incorporated member checks in various phases. During data collection, we kept in touch with two (key) respondents to clarify unclear situations. After the data collection, we presented our initial results and a first draft of the chapter to gain feedback on our analysis. The respondents recognized their work in the analysis and pointed out minor misunderstandings (e.g. datasets were updated on a weekly basis instead of a daily basis). We presented the final draft of the chapter to our respondents to gain their permission to use quotes and asked them to check our use of this material (Mortelmans, 2007).

Practice	Who	Main involvement	Data collection
Data science	Internal data scientist A	Team leader, all research projects	Interview, observations (17 days)
	Internal data scientist B	Antipsychotic project	Interview, observations (18 days)
	Internal data scientist C	Antipsychotic project	Interview
	Internal data scientist D	Antidepressant project	Interview, observations (16 days)
	External data scientist E	Antipsychotic project	Interview, observations (6 days)
	External data scientist F	Antipsychotic project	Interview, observations (7 days)
	External data scientist G	Antidepressant project	Interview, observations (11 days)
	PhD, data scientist H	Various projects	Interview, 12 days present in observations
Psychiatry	Psychiatrist I	Professor, medical head of department, initiator of the data-driven team, various research projects	Interview, observations (4 days)
	Psychiatrist J	Department head, studied schizophrenia, antipsychotic project	Observations (2 days)
	Psychiatrist K	Studied social cognition in schizophrenia, antidepressant and antipsychotic project	Observations (3 days)
	Psychiatrist L	Studied gene-environment interaction and antipsychotic project	Observations (3 days)
	Psychiatrist M	Studied developmental disorders in children. Advisor to the data-driven team	Interview
	Psychiatrist N	Studied genetic predispositions in relation to mental illnesses, antidepressant project	Interview, observations (3 days)
	Researcher O	Assistant professor, expertise in neuroimaging. Advisor to the data-driven team	Interview
	Researcher P	Assistant professor, expertise in genetic epidemiology, antipsychotic project	Observations (2 days)
	Nurse Q	Involved in antidepressant project	Interview, observations (1 day)
	Nurse R	Advisor of the data-driven team	Interview

**Table 1.** Overview of key practitioners of both communities

Results

In the next section, we will present the hospital-based data-driven initiative and the data science and psychiatry practitioners involved. Afterward, we analyze the process of trading epistemic differences, illustrated by two examples derived from data science subprojects developed in the initiative.



## The trading zone

We learned that not all the psychiatry practitioners greeted the data scientists with enthusiasm, when they arrived at the hospital. There were diverse criticisms. Some psychiatrists were skeptical because of previous costly, time-intensive ICT projects had limited added value. Others said that the rapid nature of developments in data analysis would quickly outdate any initiative. They argued it would be better to wait and see how techniques developed in other contexts. Also, some questioned the methods used. Psychiatrist M explains:

*“When they first presented the ideas [on the initiative], I really had to hear them a few times and think about them. That’s how it works for me! [...]. I’ve learned to be very suspicious of connections that you find in data. The more you look for certain connections, the greater the chance that you will find one, but it could well be complete nonsense” (Interview psychiatrist M).*

The data scientists decided to focus their attention on psychiatry practitioners who were enthusiastic about the initiative and willing to cooperate. The psychiatry practitioners who participated in the data-driven initiative hoped to find new sorts of information that could help them to improve the treatment of patients. Several psychiatrists described the unpredictability of psychiatric diseases. They mentioned, patients with long-lasting depression who unexpectedly recovered, patients who benefited from strange combinations of medicines or patients who developed severe, unanticipated side effects. Psychiatrists relied on a combination of scientific evidence and their professional experience to treat patients but realized that their experience did not always provide clear evidence and contained biases. As psychiatrist I explained:

*“We try to treat patients to the best of our ability, but often we would like to have more information. There are guidelines for extreme situations, but there is a large gray area. Those guidelines are fine, we don’t need to change them, but we do need more guidance in this gray area. We don’t know much; there is simply not a lot of evidence. We need some sort of ... intermediate evidence” (Interview psychiatrist I).*

The data scientists kept in touch with the medical practitioners through face-to-face meetings, via e-mail and by accompanying them on the daily rounds. We observed practitioners in both communities discussing the data that could be included in a prediction model, validating preliminary machine learning outcomes and brainstorming on verifying the decision-support models. We highlight two cases: the first case is illustrative of the majority of the negotiations we observed. The second case because negotiations were most fierce and challenging.

## Negotiations during the development of a prediction model for antidepressants

The first case is of a typical meeting between data scientists and psychiatrists. The meeting is part of the trajectory in which the data scientists are trying to develop a prediction model that will help determine which antidepressant is best for certain groups of patients. During this meeting, data scientists (A, D and G) and psychiatrist N are sitting in N's small office. The data scientists have extensively studied prescription data and guidelines and have had similar meetings with other psychiatry practitioners. Now, they want to validate their findings and come to a decision about how to proceed.

---

### Case 1. How to define a successful treatment?

The meeting starts with a discussion about prescription behavior and the possible side effects of antidepressants. Data scientist G explains that the data scientists are trying to distinguish between "successful" and "unsuccessful" medical treatment as the first step in developing their prediction model. After consulting with various psychiatrists and studying the guidelines, he is considering the following conditions. First, a drug must have been prescribed for a minimum amount of time so that it has enough time to be effective. Second, the Hamilton depression score [a measure of the severity of depression] must be at least 50% lower than on the patient's admission date or should be equal to 8 [indicating sufficient improvement or a "normal" value]. He is thinking of including Beck scores and the Functional Disability Inventory [both aimed at measuring depression severity]. The only problem is that they are embedded in text fields in the patient files, which makes them hard to use. Data scientist D adds that they have heard from other psychiatrists that MADRS [another depression scale] might be relevant. However, they noticed in the data that medical professionals in this hospital do not use the MADRS scale.

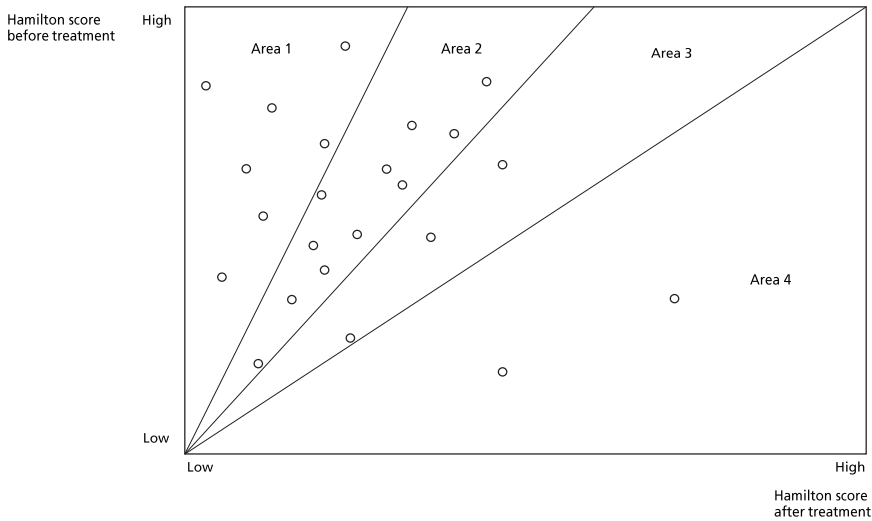
Data scientist G argues that the Hamilton score seems most important. Psychiatrist N nods in agreement, but sighs: "What a shame that it's so hard to analyze free text fields in patient files." Data scientist G explains that they are unsure about the "50% lower" or "equal to 8" criteria as these seem quite random threshold levels. Psychiatrist N disagrees because, according to him, most medication studies use the same threshold levels. He gives some context: "A Hamilton score can go up to around 25 points when someone is severely depressed. You should be more worried about measurement variation because medical students fill out the questionnaire with the patients during their internships and they don't always take it seriously." This is new to the data scientists, but, as data scientist A remarks; "We do not really have another measure that we can use, so this is our best option." Psychiatrist N agrees.

Data scientist G shows a graph of the current distribution of Hamilton scores for patients treated with antidepressants [see Fig. 1 for a similar but fictional graph]. The graph reveals that only a small portion of patients lands in the "lower or equal to a score of 8" (area 1), meaning that they have had "successful treatment" according to the current criteria. A more substantial group of patients have had successful treatment if "50% improvement" (area 2) is also considered. A few patients show a partial response (area 3), and even fewer show no or a negative response (area 4). The data scientists suggest also including patients with partial response (area 2 and 3) in their definition of "successful", as their Hamilton scores improved after medical treatment.

Psychiatrist N looks at the graph more closely and remarks that the graph is different than what he would have expected, based on the scientific literature. He asks the data scientists what sorts of data they used to make the graph. He goes to his computer, intending to look up a scientific study, when suddenly his phone rings, interrupting the meeting. Something is happening to one of his patients. It seems quite serious because psychiatrist N excuses himself and, heading off to the ward, adds, "We must continue this conversation next week." (Observation notes).

---

This first case shows how the data scientists and psychiatrist negotiate on the data to include and adequate threshold levels. During the meeting, they were able to navigate some of their differences and make the decision to base the prediction model solely on the Hamilton data. At one point, the data scientists were afraid to lose their last data source (Hamilton data) when psychiatrist N started to talk about measurement variation. At that moment, both parties decided pragmatically to use the Hamilton data as it was “the best option”.



**Figure 1.** Graph of treatment response of individual patients

The trading process took place during formal meetings and it helped that the data scientists learned to speak in the vocabulary of questionnaires and depression scales that is understandable for the psychiatrist. Simultaneously, psychiatrist N confirmed the struggles of the data scientists and contributed to the negotiations by stating “what a shame that it’s so hard to analyze free text fields”, thereby showing that he understood the importance of these data for the data scientists. The trading process was also stimulated by the systematic approach of the data scientists. This virtue was recognized by psychiatrist N. He understood the importance of determining outcome measures, as it was also an important part of the scientific research he conducted in the past. The graph also helped to visualize the threshold dilemma and thereby contributed to the cross-boundary interactions.

There were also epistemic virtues that needed to be negotiated. The data scientists looked for completeness in the data. They wanted to include as much data about depressions as possible; several measurements, qualitative data and data of patients that partially respond to the treatment. At the same time, psychiatrist N was more selective about the data to use

and threshold levels to select. N referred to scientific studies during the meeting. With his mentioning of scientific literature, he brought in authority and the rich scientific history of psychiatry. He stressed the importance of being selective in the kinds of data to include and thresholds to uphold while simultaneously highlighting the embeddedness of psychiatry in practice.

## Negotiations during the implementation of a prediction model for antipsychotics

The second case highlights a trading process about the importance of statistical significance between data science and psychiatry practitioners. The discussion became most evident during the antipsychotics subproject. Within this project, both parties had to negotiate the standards and measures to uphold for determining the performance of the prediction model and the conditions under which the model could be implemented in healthcare practice.

---

### Case 2. When to introduce a machine learning model in daily healthcare routines?

The psychiatry practitioners needed certainty that the models would significantly improve current healthcare practices. They argued that, while not perfect, statistical significance would be a necessary step to test the performance of the models before introducing them in daily healthcare practices, since it would help determine the reliability and ensure that the success of the model was not based on chance. They argued that it was necessary to set up a RCT-inspired approach, in which the performance of the model would be tested against care-as-usual.

The data scientists did not completely understand the focus on statistical significance. Some of the data scientists previously worked in the commercial sector, where they never had to work with statistical significance. They searched for findings with the biggest impact and tended to look at prediction accuracy instead of statistical significance. Data scientist G explained: "If you started to talk about significance in such a [commercial] setting, everybody would look at you and say: "what are you talking about?" (Interview data scientist G). The data scientists explained that they use their own methods to validate their findings and to see if their model is performing well:

Data scientist E: "Significance is just not that interesting with techniques like machine learning! Well, that might be a bit exaggerated, but ..."

Data scientist C: "It is not the ONLY measure!"

Data scientist E: "We often use lots of data to generate our models, when significance is not necessarily the most important measure. There are other ways to validate that your model is doing well and in what circumstances." (Interview data scientists C and E).

The notion of statistical significance caused misunderstanding and miscommunication. At the beginning of the project, the data scientists had proudly presented some of their initial findings when a psychiatry practitioner asked if the results were statistically significant. The data scientists answered that such a statistical test was less important nowadays. It was more important to validate your findings with medical professionals. Psychiatric researcher O was critical about this approach:

---

---

"I don't agree with the approach. I mean, I think you still have to make sure that a result is statistically significant and not produced by chance, so I wonder about dismissing this statistical test as somewhat unimportant." (Interview psychiatric researcher O).

The presentation of the data scientists had ended with a fierce discussion and became a story that was retold in both communities.

During the end of our fieldwork, uncertainty remained about the conditions under which the data scientists could implement their machine learning models in practice. The psychiatry practitioners continued to hold the opinion that statistical significance was crucial for the evaluation of the prediction models. Some argued that data scientists should not even present non-statistically significant findings to them, as this could negatively influence their behavior. Data scientists stated, by contrast, that "it would be a pity not to use all the non-significant results. You waste a lot of information that could be used to our advantage" (data scientist F).

The data scientists tried to obtain some leeway by organizing meetings with psychiatry practitioners that were interested in their methods. This allowed the data scientists to explain their alternative validation methods. Also, strategic meetings with more senior staff were organized to bring both parties together and make workable agreements. Simultaneously, the data scientists agreed with the RCT-approach to compare their models with care-as-usual as they understood the benefit of such a comparison.

---

This second case shows how data science and psychiatry practitioners negotiated about the implementation of the machine learning models in practice. There was a partial compromise as both parties agreed on the RCT-approach to compare the prediction model against the care-as-usual. The data scientists did see the value of this approach and understood the responsibilities of the medical practitioners. However, the negotiations were far more intense and developed differently than in the first case.

In this case, trading was for the large part done by backstage politics. The data scientists did set up meetings with interested medical practitioners to make them familiar with their epistemic culture and to teach them their "language". Besides, the strategic meetings with senior staff were organized. These persons were respected by both communities and could stimulate the negotiations. Also, it helped that both the data science and psychiatry practitioners felt that data, models and outcomes had to be verified and recognized the importance of having validation standards as a virtue. The negotiation process that unfolded was about the sort of validation process to use.

There were epistemic virtues that needed to be negotiated and this was not straightforward. The psychiatry practitioners, while enthusiastic about the prediction models, were reluctant to introduce these models into their care routines. They needed certainty about the functioning of the model and the methodologies used. This was not necessarily unwillingness on their part, as it also ties closely to the various epistemic responsibilities and risks that they face (e.g. treatment responsibility and malpractice claims). The data scientists could be more flexible in the outcomes and evidence they selected as they do not face these responsibilities and risks. The flexibility was also visible in the statement of data scientist E when he talked about validating that "the model is doing well in certain

circumstances". This shows that, for the data scientists, the necessary degree of flexibility or certainty was dependent on the impact of the analysis. This would mean that a data-driven model for deciding on deep brain stimulation needed more assurances than a model used for scheduling staff. This more dynamic use of information meant that data scientists often accepted more flexibility in their methods and stressed that they might have to revise their analyses after a few years or based on new data.

The trading process, in this case, was more difficult than in the first case. It was not helpful that both communities retold stories (such as about the discussion on statistical significance) within their own communities as it emphasized the differences between the communities. The negotiations were already difficult, especially because the virtues (flexibility versus certainty) were so different and complex statistical theories clouded the discussions.

The field of psychiatry is used to be working with causal inference and research that looks for causation. This type of research aims to give answers to questions such as: "how much more effective is antipsychotic A compared to antipsychotic B for (a group of) patients?". Medical practitioners can use this information to change the treatment practice for their patients. The practice of prediction modeling of the data scientists is based on another statistical theory. This type of research tries to predict what the chance is that someone improves when prescribed a particular type of antipsychotic. As such, it allows medical practitioners to be reactive to situations. It, for example, presents the chance that a side effect will occur in a patient and this enables psychiatrists to anticipate on the side effects. The discussions about significance illustrate that it is complicated to understand and act upon these different sorts of information. Causality thinking forms such a considerable aspect of the culture of psychiatry and is so embodied by all practitioners that might be difficult to grasp alternative approaches and vice versa.

## Discussion

In this chapter, we study the convergence and divergence of the epistemic cultures of psychiatry and data science. We argue that such similarities or differences are not pre-given but depend on situated negotiations and actions. Through the notions of "trading zones" and "epistemic virtues", we have been able to analyze how differences between epistemic cultures are negotiated in situ by data science and psychiatry practitioners. In this discussion, we argue that our combined theoretical framework offers a fresh way to study how cooperation between diverse practitioners develops and where it can be improved. We make a call to bring epistemic differences into the open as it makes a grounded discussion

about the added value of data-driven initiatives and the role they can play in healthcare, possible.

Our ethnographic approach showed that practitioners from both communities actively sought collaboration to improve care for psychiatric patients. Many epistemic differences needed to be negotiated. Situated judgments about the need for being complete or selective in the data to include and about the amount of flexibility or certainty that is permissible, revealed the role of specific epistemic virtues. We found that it is not simply the case that diverse communities persistently try to incorporate the epistemic virtues of the other (Galison, 1997: 401, 807). The dynamics are more complex and far messier as epistemic virtues get their meaning in concrete practices.

Our study of the data-driven initiative shows how epistemic differences are traded locally and temporarily. Our work aligns with earlier work on trading zones in highlighting the importance of boundary objects (the graph in the first case), interactional expertise (data scientists learning the medical language in the first case) and socialization (data scientists becoming embedded in the culture of psychiatric measurements in the first case) in stimulating cross-boundary interactions (Galison and Stump, 1995; Galison, 1997; Knorr Cetina, 1999). Our research adds to this literature by highlighting two additional processes. First, our work shows the importance of validation of the other culture in helping to navigate differences. The remark of the psychiatrist in case 1 about how unfortunate it was that free text fields are so hard to analyze, showed that he learned enough about the culture of the data scientists, and by explicating these differences he showed respect for the other culture. Second, shared epistemic virtues, such as the systematic approach and the validation standards, helped to stimulate the negotiation processes and create space for negotiation, despite differences.

In both cases, the communities were able to navigate some of the differences in culture. However, the negotiations in the second case were more complicated, because epistemic virtues were too different (e.g. flexibility and certainty) and hidden within complex statistical theories. What also complicated the interactions, was the retelling of stories that emphasized the epistemic differences.

The study has two limitations. First, our distinction between the two subcultures might be considered too simplistic or artificial as there is never just one single data-driven culture nor only one medical culture (Harmen and Galison, 2008; Knorr Cetina, 1981). We argue that while there is variation, there is evidence that shows an overarching medical and data-driven epistemic culture as we described in our theoretical framework. For the purpose of this study, aimed at exploring the similarities and differences in epistemic culture, we drew

a division based on the participants' positions in the hospital-based initiative, but different divisions could be made.

Second, while rich in detail and extended with other data collection methods, our ethnographic study of this data-driven initiative lasted six months. The relatively short timeframe made it impossible to observe, more longitudinal process changes, such as the development of shared language or more gradual changes in epistemic virtues. As such, more longitudinal ethnographic research projects, following data-driven initiatives over an extended period, would be a welcome addition.

We argue that it is important to bring epistemic differences into the open as this enables a grounded discussion about the added value of data-driven initiatives and the role they can play in healthcare. This study shows the added value of ethnographic, action-based research in making subtle differences in epistemic cultures visible in concrete data-driven initiatives. Data-science initiatives are not impossible or simply possible in healthcare and epistemic virtues help to study where problems arise and how practitioners deal with them. This approach helps to make a more grounded discussion possible, about the type of data to include, thresholds to uphold and the necessary steps for the implementation of prediction models.

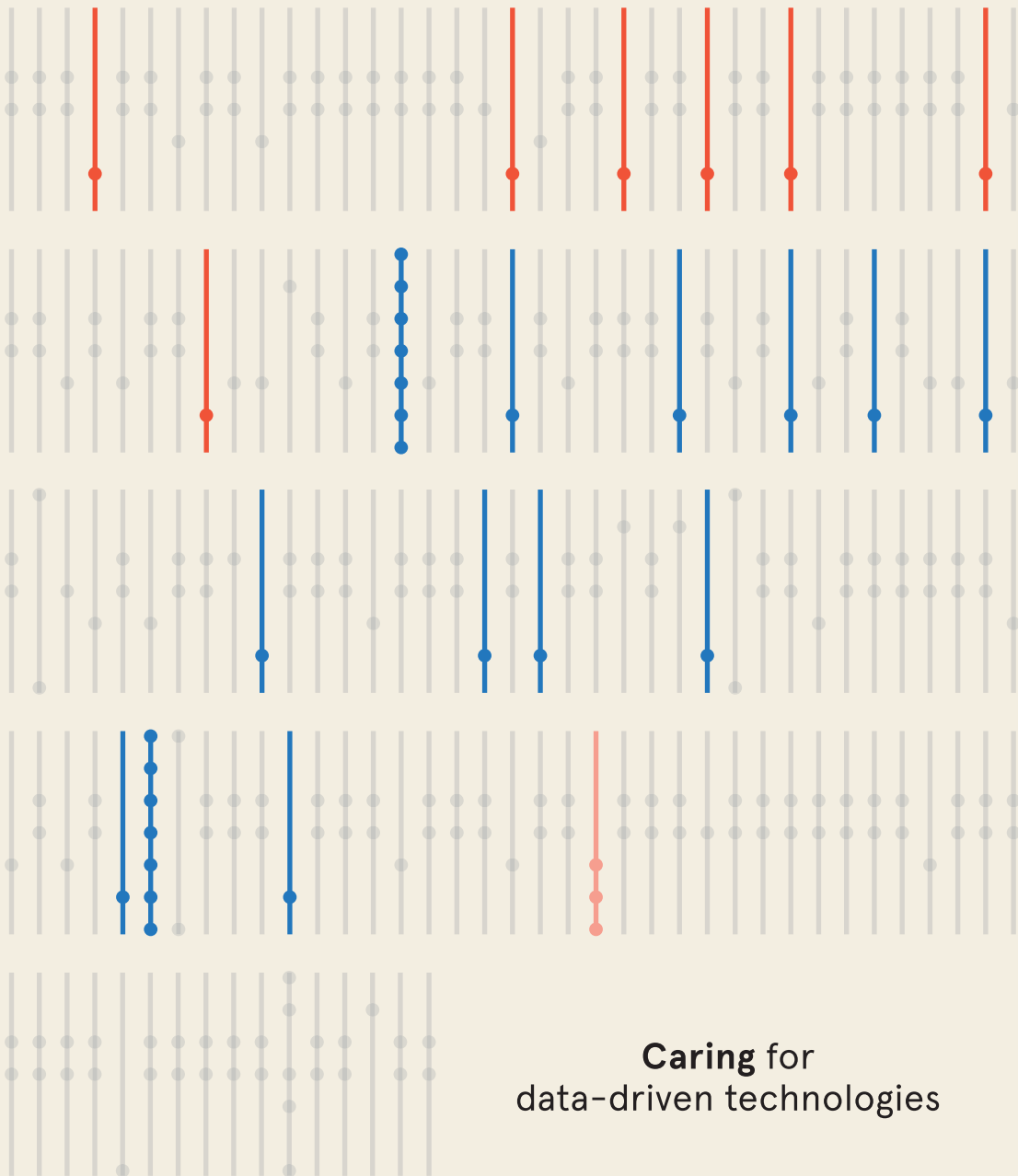
Such ethnographic studies can subsequently be used to organize moments of reflexivity for data-science and healthcare practitioners, for example, by starting initiatives and experiments that stimulate discussions. Think of organizing focus groups and data deliberation sessions, involving the relevant communities, such as professionals, patients, data scientists, researchers and policymakers, to explore the epistemic differences and similarities (Haan et al., 2018; Madsen and Munk, 2019; Moats and Seaver, 2019; Ziewitz, 2017). Epistemic cultures are not fixed, and they can change. The epistemic cultures of data science and psychiatry are in some respects relatively similar but also differ in myriad ways. It is highly relevant to study the differences and see how both cultures can learn from each other.





Under review at Science Technology, & Human Values as:  
Stevens M and Beaulieu A – Relating to epistemic responsibility:  
the machine learning entanglements of ten mental healthcare  
organizations in the Netherlands.

# Chapter 6





## Introduction

*The teacher stands in front of the fourteen healthcare professionals in a small conference room during the first day of the machine learning course. He explains a dilemma that the professionals could face with their new machine learning prediction models: “what if you have calculated the chance that a patient will recover, but you know that there is only a small chance. How do you present the outcomes of your machine learning model to other healthcare professionals? Do you not give them a new problem instead of helping them? We face many challenges with these new techniques that are different from the types of challenges encountered when doing more traditional effectiveness research” (Observation day 1).*

The introduction of machine learning is not without dilemmas and uncertainties: what are its many consequences and for whom? What is it about the new approaches that make us feel like they are a game-changer in terms of responsibility? These are the larger questions we seek to address via an ethnographic exploration of projects that seek to incorporate machine learning in mental healthcare. We use the lens of care (Puig de la Bellacasa, 2017) to reframe responsibility as an effect of the entanglements of humans and technology. Rather than seeing responsibility as something to be assigned to a specific party, we consider how it is produced over time and in interactions between actors. This reframing opens up new possibilities to address pleas for responsible research and innovation (e.g. Zook et al., 2017).

Machine learning approaches are increasingly used in healthcare to handle patient data, treat chronic diseases and develop medical procedures (e.g. Dwyer et al., 2018; Jiang et al., 2017). In the case of machine learning, this means that algorithms serve to find novel patterns in data that can be used for making predictions about the data. Machine learning includes a range of techniques that involve different combinations of human and automated processes, from the human labeling of certain inputs (supervised learning) to approaches without human supervision (unsupervised learning). Across these approaches, machine learning marks a shift in focus in statistical techniques for processing data, since it is less oriented to causal investigation and hypothesis testing, and more focused on prediction and finding correlations.

Much of the hope attached to machine learning is based on assumptions about the relationship between technology and epistemology: these new technologies drive new insights and information (Stevens et al., 2018, Beaulieu, 2001). This assumed relation also shapes fears about algorithms as they are seen as opaque, increasingly self-learning and somehow out of reach (Ziewitz, 2017). This leads to uncertainty about the responsible use

of algorithms and their outcomes, to which we are increasingly confronted (O'Neil, 2016). This is not due to a mechanistic impact that an external, novel technology might have. This is the case because machine learning approaches shape the epistemic culture of the medical community, insofar as they change the relation to the object of knowledge, and to the reasoning and evidence that determine what and how we know (Knorr Cetina, 1999; Stevens et al., 2020; Wouters and Beaulieu, 2006).

In this chapter, we specifically focus on epistemic responsibilities - the responsibilities in the process of knowing (Simon, 2015)<sup>1</sup>. Traditionally, responsibility is seen as a state or moral value that is available to individual actors - actors that are sovereign, rational and autonomous, and who take responsibility intentionally (see Kroes and Verbeek, 2014). These studies focus on ways in which a medical doctor takes responsibility, for example, by selecting certain types of information using complex reasoning styles, while also demanding that certain epistemic criteria are met (Van Baalen, 2019).

With the development and uptake of algorithms and machine learning trajectories, scholars have posited that responsibility is in crisis, given the increasing interconnections of humans and technological systems and the growing mediation of sources of information (Mittelstadt et al., 2016; Simon, 2015). First, machine learning trajectories are created in highly entangled and complex environments consisting of humans and non-human agents, such as databases, technologies, infrastructures and organizations. In these environments, technologies, knowledge, knowers are engaged in complex relationships, see their power shifting, and find themselves to be interdependent rather than hierarchically organized (cf. Callon and Law, 1997; Latour, 1987). In these complex networks of entanglements, the traditional, individualized conceptions of epistemic responsibility are increasingly difficult to uphold and need rethinking (Mittelstadt et al., 2016; Simon, 2015).

A second and related point is that there is a growing mediation of sources of information (Mittelstadt et al., 2016; Simon, 2015). Our ways of knowing have always been enmeshed with material and often technological objects (Beaulieu et al., 2014; Verbeek, 2015). We know because we can measure temperature with a thermometer, see bacteria under microscopes and can engage with bodies through X-rays and CT-scans. Our relationships to material and technical objects are increasingly informational so that we interact with layer upon layer of mediation. Different "black boxes" are created and this leads to the feeling that we do not know where information comes from nor how it is produced.

Rather than seek to peel back the layers to a solid set of clearly defined objects, essential roles, hard criteria and fundamental responsibilities, we set out to embrace the complexity of these relationships as a starting point. Instead of centralizing epistemic agents, we

focus on the entanglements to explore how the growing connections generate epistemic responsibility.

## **Responsibility-in-the-making in machine learning trajectories**

Our argument builds on two elements from STS and feminist literature that enable us to take a novel approach to epistemic responsibility in machine learning trajectories. STS has taught us a lot about how to interrogate human decision-making systems and knowledge-systems (e.g. Knorr Cetina, 1999; Latour, 1979; Timmermans and Berg, 2010). Feminist analyses of science have focused on the importance of the relational nature of knowledge and politics (e.g. Barad, 1996; Puig de la Bellacasa, 2017; Suchmann, 2009; Tronto, 1987). Taken together, this means that first, we do not see growing entanglements as problematic and standing in the way of responsibility, since our feminist/STS-informed starting point means that responsibility precludes a sovereign, rational, reason-giving actor. Rather, we see these impure entanglements as generating responsibility and set out to discover the sort of responsibility they produce. A second important element of our approach is that by seeing relations as performative, they are therefore external neither to knowledge nor to responsibility. The result is that ethical obligations associated with knowing shift from being driven by moral principles or shaped by external technological forms, to being embedded in material forces that are involved in the constraints for the everyday continuation and maintenance of life (Puig de la Bellacasa, 2017: 22).

For this study, we choose an ethnographic approach to embrace all the nuances, efforts and workaround in practice. It provides a detailed empirical understanding of how responsibility is taken in real-life machine learning initiatives. In the middle of all the hype and promises that accompany machine learning, it is sometimes forgotten that these machine learning trajectories are not only an abstracting, formalizing tool but also become part of our social life because of the sites and interactions where they are implemented (e.g. Stevens et al., 2020). This means that a potential deficit of responsibility is not an a priori feature of using machine learning, and that such a lack of responsibility is rather the outcome of specific entanglements.

The site at the center of our ethnographic exploration is one in which machine learning becomes entangled with healthcare practices. As we engaged with our field site, we came to focus on moments where responsibility was being put into question or was being put to work, and interactions where responsibility was reeled in or delegated.

The site arose from a mission to enact a responsible approach to machine learning: frustrated by the empty promises, obscurity and complexity that surrounded machine learning, a four-month machine learning course for healthcare professionals was initiated by an expert at a Dutch non-profit research institution that focuses on the mental healthcare sector. Fourteen healthcare professionals were selected to learn the basics of (supervised) machine learning, while also being mandated by their institutions to pursue a project that could be significant for their organization.

The course was an opportunity for us as researchers to engage with a complex, emerging responsibility, since learning foregrounds and backgrounds various relations and brings responsibility-in-the-making into relief. The field site was full of situations where we could join these professionals as they asked, in different ways, “how to care” for information and knowledge in the healthcare sector. When this question is insistent but not easily answerable, possibilities for imagining different futures open up (Puig de la Bellacasa, 2017: 7). This is what we saw happening around epistemic responsibility. As the group of healthcare professionals was taught about machine learning, they continuously contrasted machine learning with approaches more familiar to them, such as randomized controlled trials, where the role and responsibilities are more settled into familiar patterns. The various relations needed to be able to engage with machine learning were also taking form as the professionals set up their own projects in their institutions. The same applies to the algorithms: they were still partly open and being developed.

Several features make this healthcare setting especially relevant to study epistemic responsibility. First, the healthcare sector traditionally has a strong epistemic culture that favors high-quality evidence to guide treatment decisions and organize the field. This also means that the field has a tradition of careful weighing of new (epistemic) practices. Furthermore, the mental healthcare sector is an interesting case to study the introduction of data-driven initiatives, since this field is characterized by considerable uncertainty relating to disease ontology and treatment effects. Moreover, it relies primarily on narratives of patients and qualitative questionnaires to make sense of the patients’ conditions and to guide treatment decisions. This means that epistemic cultures of the mental health care field have a relation to their object that emphasizes context and particularity of cases - rather than abstraction and the search for formal patterns typical of machine learning applications. In this specific sense, the struggles of actors to assign meaning to algorithms contrasts with fields where automation and datafication have a much longer history, for example, genetics research (Lee and Hegelsson, 2020).



## Epistemic responsibility

Epistemic responsibility is used in philosophy to describe duties in the process of knowing, namely giving and accepting reasons for accepting or believing in what is being stated or not (Simon, 2015: 157). Central to this work is the presumption that responsibility is understood and analyzed in terms of sovereign, rational and autonomous humans that acts out of free will and intentionally (Nissenbaum, 1998).

As described in the introduction, it is often argued that recent developments in the field of (informational) technologies have made the process of knowing and taking responsibility increasingly difficult, due to the growing interconnections of humans and technological systems and the intensification in the mediation of sources of information (Mittelstadt et al., 2016; Simon, 2015). This leads to fears, uncertainty and reduced feeling of personal responsibility, which could lead to an increase in the execution of otherwise unjustifiable actions (Mittelstadt et al., 2016).

Scholars - primarily within the field of philosophy of technology – have rethought the conditions for being responsible (Kroes and Verbeek, 2014). Traditionally, moral agency is necessary for responsibility. Since moral agency depends on intentionality and freedom, this approach limits moral agency (and thus responsibility) to human beings. And while this traditional notion of responsibility does not easily account for collective actions and networks, this position is still defended by many ethics' scholars.<sup>2</sup>

A contrasting approach is taken by Floridi and Sanders (2004) as they work on disentangling moral agency and moral responsibility. They argue that technologies can be involved in moral situations, but since they do possess neither freedom nor intentionality, they do not have moral agency and cannot be responsible. Instead, they introduce the concept of "mind-less morality" for non-human actors that act autonomously, interactively and adaptably, such as advanced artificial intelligence systems. This additional category means that, while non-human actors cannot be held responsible since responsibility requires intentionality, they can be held accountable. This diversifies the kind of moral roles that technologies can play while maintaining responsibility as a human attribute.

However, by assigning roles and responsibilities, and taking the autonomous individual as a starting point, these approaches start and keep us in a very rational and structured view of knowledge and responsibility. They tend to over-value autonomy and agency, thereby devaluing other relations (such as human-technology, or symbolic-material) and reinforcing asymmetries between humans and non-humans (see Verbeek, 2011; 2014 for similar

critiques). These critiques are also well-known within STS work that foreground diverse relations (cf. Callon and Law, 1997; Latour, 1987).

Others in the field of philosophy of technology recognize the close intertwinement between humans and non-human actors and describe networks in which technologies have some form of moral agency in addition to humans (e.g. Introna, 2014; Verbeek, 2014). These scholars argue that conditions for moral agency (intentionality and freedom) are less strictly reserved for humans than classical ethical theories assume. For instance, in the work of Verbeek (2014), intentionality -interpreted as “giving direction to actions”- is also recognizable in technologies, and freedom -conceived as “the ability to relate to what influences you”- takes shape in the interaction between humans and technology. This does not mean that technologies have moral agency just like humans do, but rather that we should see moral agency as the outcome of complex interactions between humans and non-humans (e.g. Verbeek, 2014: 80).

There are critiques on these “relational” approaches. First, within critical algorithmic studies and STS in general, there is critique on studies that highlight relationality as they are often too apolitical and blind to the power struggles of weaker actors as well as to political effects (Giraud, 2019; Lee and Bjorklund Larsen, 2019). Second, a critique from philosophy of technology is that studies that highlight the relationality of responsibility, often remain unclear and a-specific (Simon, 2015).

To address both critiques, we introduce more recent feminist/STS scholarship on care in this chapter that focuses on relationships, knowledge and power. These scholars insist on commitment to ongoing relationships in knowing, with the implication that responsibility is an outcome rather than a fixed position (Suchman, 2009). This body of work pushes ideas about moral agency as the outcome of complex interactions between humans and non-humans further by linking epistemology and ethics to “care” and arguing for studies that show the complexities, interdependencies and tensions involved. This combination allows to foreground the labor, affect and ethical-political dimensions of taking responsibility in networks (Puig de la Bellacasa, 2017: 5), enabling more attention to political and power struggles. It also encourages an ethnographic approach that focuses on concrete interactions in real-life situations and that supports the reflexive taking of responsibility.

## Care approach

Recent work on care offers a way to think through human and non-human entanglements. The care approach moves away from the direction of the philosophy of technology where

scholars are pushing for sharper delineation and hierarchy of agencies (some being “only accountable” and others “fully responsible”). Instead, this approach makes it possible to take a different starting point that is not based on anxiety to be alleviated with control strategies.

Care highlights the doing and the active making of the world while being connected to others (Puig de la Bellacasa, 2017). Puig de la Bellacasa (2017) gives an example of how worms take care of our soil. They provide care that is essential for our survival, but worms do not consciously engage in it. However, there is a lot to be gained to study how worms care for us and what they need to keep doing for this caring. This draws attention to how we need to care for the worms ourselves (e.g. by not making the ground too dry or empty of nutrients, not to harden it by using heavy machinery). In this example, knowledge of what good soil is, is gained by understanding the connections in a web of relations that make up soil. Taking responsibility means acknowledging the care work needed on the part of different entities - humans, worms and soil. Puig recommends that questions of responsibility be explored pragmatically and creatively, a project that may be most feasible in sites of entanglement (her case is permaculture; animal care, conservation and activism are other promising ones), where responsibility is tied to doings that are essential for prolonging life.

Similarly, in our case medical professionals, (data) scientists, machine learning algorithms and data registries intimately work together in machine learning trajectories. Together, they work to select, rework and process data. We study these entanglements and work with the care perspective to highlight what human and non-human actors undertake to care for the data, methods, results and how, through this caring, they take responsibility (Puig de la Bellacasa, 2017).

The care perspective highlights the labor performed by actors in the network (Puig de la Bellacasa, 2017): the doing and the active responsibility-in-the-making while being connected to others. Care should be understood analytically, instead of merely as a well-intentioned, good attitude (Tronto, 1987; Puig de la Bellacasa, 2017).<sup>3</sup> The care approach is thus also about maintenance: if the caring stops, things fall apart.

The approach entails a committed, speculative sensitivity to human and non-human actors less able to make their voice heard or concerns known. A care approach attunes us to the care-like work done by non-human others (such as machine learning algorithms) in sustaining the everyday while prompting us to pay closer attention to the particulars of their existence. So, it is worthwhile to imagine non-human others (such as algorithms and computers) as caring.

Thinking with care is also a way to address power and politics: as we will see, there are struggles; to decide what to care about, uncertainty about how to care for algorithms and patients (apart or, more interestingly, in conjunction) and frustrations because the algorithm may or may not care about certain things.

To summarize, we could contrast the value of our approach as the difference between asking what is the situation, such that it has responsibility as an effect, rather than who is the responsible actor. We opt for the first, situated approach that is shaped by the scholarship on care. It foregrounds the relationality of knowledge processes, so that: (1) entanglements are no longer seen as problematic and standing in the way of responsibility, and so that the connections become both (2) worthy of inquiry and (3) generative of responsibility.

## The case

*"I walk up the stairs to the conference room for the second day of the machine learning course. A couple of the participants are already sitting at the long table and I take a seat. Once again, I notice that there is a shortage of extension cords and electric sockets. Everybody wants to plug in their laptops to make notes and prepare for the R practicum later today. Our teacher, Z, brings in the janitor and he is so kind as to start looking for a few extra extension cords. I decide to use this opportunity to grab a cup of tea in the hallway. Once back, there is some somewhat strange, dark music playing. I ask medical professional B what is going on. B explains that this particular song is based on a machine learning algorithm and says that "Z told us that this is a quiz, we have to guess which algorithm it is!". This confuses me: do we really have to guess this? I have no idea where to start! But then I realize that it is only a joke" (Observations day 2).*

This vignette contains several figures that are important for our analysis. First, there is care work going on. There is a need to arrange connections; laptops will be essential to the work to be done, and humans in various roles (participant, researcher, teacher, janitor) ensure that the technologies are properly cared for so that these will later enable the humans to engage with other technologies, such as the R software platform and the algorithms that will run on the laptops to make predictions based on their data. A second figure is the reconfiguration of relationships to algorithms: the music provides a sensory invitation to engage with novel objects. The joke foregrounds whether it is even possible to guess the identity of an algorithm from its musical performance. Humor contains a call to consider what kind of epistemic work can be done, and plays with the displacement of the boundary between the serious and the fanciful, between what can and cannot be

known. While clearly an exaggeration of the possibilities, this joke articulates the invitation to the participants of the course to open themselves to knowing differently, to the logic of correlation and pattern-making that characterizes algorithmic knowing.

\*\*\*

The machine learning educational course for medical professionals, the starting point for our site, was initiated in 2018 by Z of a Dutch non-profit institute to give medical professionals insights into the many possibilities of machine learning algorithms. This institute specializes in mental health research and takes an active role in the distribution of (innovative) research and technologies in the Dutch mental healthcare field. Frustrated by the empty promises that surrounded machine learning, he aimed to train professionals the basics of supervised machine learning, while supporting them in performing their own machine learning research project that could be relevant for their organization.

The course was advertised in online newsletters and not long after, many healthcare professionals and organizations showed their interest. Healthcare organizations and professionals recognized the importance of the new statistical techniques for their organizations but often lacked (internal) expertise. The course was seen as a way to obtain this expertise, with relatively low time and cost investment.

The small-scale course lasted four months and was organized around four training days (March - June 2019). Z selected fourteen participants (A-N) working on ten projects since he wanted to make sure that there was sufficient attention for all. He was also concerned that the projects should be feasible, with regards to the skills of the participants, and in terms of the amount of time that they could dedicate to the training. Finally, he wanted to see a diversity of institutions and various types of projects included. Most participants worked for large organizations that specialize in the treatment of patients with diverse mental illnesses and/or addictions. Only participant D worked for an organization that monitors and promotes health for the inhabitants of a large Dutch municipality.

The participants used the projects to address some of the persistent problems within their organizations. Some projects were aimed at improving the healthcare organization itself, for example, by predicting treatment costs and time. These predictions could be helpful for organizations in the negotiations of care budgets with their municipalities (a necessary procedure in the Dutch healthcare system) and with the allocation of staff. Other projects were more directly aimed at improving the treatment of clients. A better estimation of relapses, incidents, referrals and drop-outs could, for example, be used to place clients sooner in the "right" treatment trajectory, to give better support and to prevent crises. The professionals hoped that prediction models could help to quicker and better recovery of their clients and help alleviate the extensive waiting lists for mental healthcare.

We reached out to Z and we were allowed to observe the four training days.<sup>4</sup> MS took field notes to capture the setting, interactions and actions during the course and paid attention to questions that the participants asked, moments of uncertainty and the discussions that were held. During the training days, Z gave lectures on the basic principles and underlying theories of machine learning. In addition to the lectures, there were R assignments and short presentations about the various projects. This allowed participants to exchange experiences. Informal conversations, questionnaires and interviews<sup>4</sup> with the participants established further relations with the participants, as did a presentation to the course participants about our fieldwork. All these interactions provided insights into the entanglements of machine learning, participants, institutions, and knowledge.

### Entanglements: a growing sense of attachment

The healthcare professionals entered the course with curiosity about machine learning and the many promises surrounding it. Most of them had an existing affective relation to quantitative research, as data analysis was already part of their healthcare practice. Yet, what machine learning would look like and how it was done was unknown to the professionals, as also explained by professional G:

*"I have been interested in new techniques for data analysis for quite some time now. For example, I subscribed to Datacamp to get to know both R and machine learning techniques. There is a lot of data within my organization, but I have the feeling that we are just scratching the surface with the traditional techniques. There is more potential in it, but my knowledge is insufficient right now" (Professional G - Questionnaire).*

G expressed that she wants to gain a deeper understanding of machine learning and the possibilities of re-using data within her organization. Many professionals used similar arguments in articulating their desire to learn more about machine learning. The professionals often referred to the hype surrounding machine learning and described how "others" started a "shallow" relationship with the techniques, while they realized that fulfilling the potential of machine learning required deeper and more targeted engagements.

In the eyes of the professionals: machine learning could help to gain new insights and improve healthcare practice. It could intensify relationships with data that is not used enough, but that is valuable in the eyes of the professionals. For example, the professionals felt that "Routine Outcome Measurement" data should be used more often to monitor the progress of their clients. In addition, machine learning could produce new insights as it helped to articulate how a consultation, as a single instance, can be seen in the context of a pattern emerging from data. By entangling individual consultations with other

data sources, the professionals could come to know more about and improve healthcare practice.

From a different angle, we can describe machine learning and its algorithms as quite distant from these professionals, and effectively disconnected from these pressing political and moral concerns. Nevertheless, the structures of these informational objects made them amenable to connecting to the professionals, for example, when they engaged them via programming language R.

The formation of new relationships between the professionals and machine learning was not straightforward. In the beginning, this new approach to data and R seemed to be more effortful than other practices more familiar to the professionals (such as effectiveness research and SPSS). The professionals got to know machine learning as needy (needing a lot of data) and difficult (giving all these errors). Data also seemed hard to access, requiring the professionals to establish closer relationships with new people and IT departments within their organizations. While simultaneously, Machine learning was subjected to what the professionals were attempting, yet was often unable to understand their syntaxes or could not make sense out of the data that was given.

As the course progressed, the professionals became more attached to machine learning and some of its logic. They invested personal time and energy, learned to speak a new language (in R) and explored the boundaries of machine learning (what it could and could not do). As they watched machine learning create new entanglements in their data, patients, therapies, and organizations, they became increasingly invested in the possibilities of machine learning, with mutual adjustments of interrelated questions, tools and data.

By the end of the course, most projects were still in progress. In contrast with the popular imaginaries of machine learning projects, there was no instantaneous mechanistic processing of pre-existing data through the deployment of “an algorithm”. What we witnessed was the nurturing of connections in complex environments, while responsibilities for knowing were slowly being redistributed and reshaped.

## **Struggles and reflexivity**

The professionals were dependent on machine learning not only because of the course and of their organizational projects, but also because of the expectations of their organizations, their own work and overarching imaginaries about the future of the mental healthcare field. They wanted to care for machine learning, in order to get new insights and improve healthcare practice, yet this led to struggles and moments of reflexivity. The professionals

felt uncertain about how to care for machine learning, its algorithms and the data, and they often did not know what they should and should not do.

Designing the machine learning model meant making assumptions about what matters and should be included. Having to make such decisions did not always sit comfortably with the professionals, because it made their responsibility in establishing a relation to the object and in shaping analysis explicit. Taking responsibility in this kind of novel situation made the professionals insecure: they often compared machine learning to techniques they were already familiar with, where what matters was already familiar to them. For example, during the second training day, feature engineering was discussed as part of the machine learning process. This meant that the professionals needed to throw out irrelevant variables and, if necessary, reformulate certain variables. This process is a way of understanding, through iteration, which features machine learning cares about. However, the participants got confused and kept asking for general rules that they could apply:

*“Professional F: “Is better to be strict or more flexible with regards to feature engineering?” Teacher Z: “I find it difficult to advise in general terms. I would say that if there is little coherence, you can probably throw the variable out, but it is a balance.” F: “But do I have to select the appropriate level for each of the variables? I find that very difficult to decide as a person; that is what entire studies are about!” Z: “Yes, that is possible” (Observations day 2).*

The professionals have been used to other approaches where they are cutting the network using (scientific) standards or customs. However, in this situation, the standards and customs have not yet been backgrounded into the everyday. This means that the change in practices and friction between business-as-usual and machine learning approaches highlighted how ethically loaded roles and responsibilities were everywhere. This breach of convention made the professionals realize that they were non-innocent participants, and made them start looking for the “general rules” and “principles” to re-externalize these decisions. However, these could often not be given in abstraction from the particularity of a case. In other words, “what should I do?” was not a useful question. In its place, professionals were led to understand that they had to do the best they could with the question, “what can I do in this context, with this data and choosing this algorithm and given what I’d like to know”. Yet, this was often a struggle and required further engagement on part of the professionals.

### **Situating and inclusion**

As the course proceeded, the professionals did engage, and they became increasingly invested with machine learning, even though machine learning was in ways different from



the professionals' conventional approaches, and even though general rules and principles for designing machine learning were often not available. This meant, for instance, that the professionals did gain an understanding of variables to keep, completing the feature engineering step. They increasingly understood how to situate their analysis and include other professionals, thereby enabling machine learning to care for their organizational problems.

The professionals experienced that this new approach to knowing required making decision and connecting to machine learning for particular situations and questions. This meant that they had to make decisions about which method, data, algorithm to select in every situation and create an iterative learning process. Teacher Z shaped the responsibility of the professionals as he continually tried to make the relations as precise as possible (what machine learning is and is not, can and cannot do). He explained, for instance, that machine learning is excellent for finding predictive relations, while randomized controlled trials are better for finding causal relationships between variables.

Simultaneously, machine learning also shaped the responsibilities of the professionals by showing its limitations. Machine learning could be trustworthy and dependable because of its mechanistic reliability in following the given commands, but it did so rigidly, not making any exceptions to its rules. This consistency helped the professionals to understand what machine learning needed and how it operated. In some specific instances, the professionals even came to trust machine learning. They learned, for example, that Lasso - a machine learning algorithm "wrapper" that can be used in R to support the analysis - is also useful for reducing their variables as it has a built-in feature selection property (Observation day 3). This means that Lasso examines what features are predictive enough to include in the model and is designed in such a way that if features do not contribute sufficiently, they will be kept out of the model. These insights led to a renewed appreciation for machine learning on the part of the professionals.

The professionals also learned that this new approach to knowing required ongoing engagement and the inclusion of people from their organizations. This meant that they needed to rely on others, ask for help from "domain experts" and together reach consensus about goals, definitions and trade-offs to be made. Involving "domain experts", meant that the professionals established sounding boards, organized meetings within their organization, started chat groups with psychiatrists and made sure they were invited to relevant meetings within their organizations. This inclusion of all the experts was especially important as the professionals realized that they needed the support of their colleagues to introduce their prediction models to healthcare practice eventually. This led to a further

extension of the range of positions to where responsibility lies and how decisions about machine learning were made

## Representing change

What we witnessed was that responsibilities for knowing were slowly being redistributed (to other actors and to machine learning itself) and reshaped (away from existing standards and customs and becoming more situated). The professionals who followed the course slowly came to understand the changes that machine learning brought, yet simultaneously they struggled with presenting and representing the newly learned lessons.

As machine learning and the professionals became increasingly invested, a mutual adjustment of knowledge, technology and ethics took place. This became visible during the second training day when there was attention for the trade-off between bias and variance. The professionals had to make this trade-off so that their final machine learning models would perform better and not follow the training dataset too little or too much, in order to avoid what is called under- or overfitting. This was new to the professionals, as also teacher Z explained:

*“Teacher Z: “When I was trained [in “traditional” statistics], I learned that I should absolutely not accept any bias. Now, I must allow a little bit of bias if it reduces variance. That is different about the types of problems tackled by machine learning” (Observations day 2).*

The language here is ethical (moral even), as it highlights the responsibilities of the professionals. In this case: the professionals should accept more bias as this results in better machine learning models. Thereby, showing that it might be possible to have slightly different configurations of knowledge, technology and ethics, depending on the situation. By accepting bias, the professionals allow machine learning to care for the data in a better way that reduces variance.

The mutual adjustments led to new sorts of freedoms and new possibilities to produce knowledge opened up. Machine learning enabled professionals to become more creative and explore new relationships as with machine learning, the professionals could test the predictive capacities of their models. This meant that they could try out different (combinations of) variables and data and see if the predictive value improved. Or in the words of professional G:

*“Professional G: “I like about this new technique that it makes me more creative than the traditional way of doing science. Now I can think much more creatively:*

*what if I now connect this with this. It unleashes many thought processes. [...] That gave me a lot of energy" (Observation day 4).*

This means that even before the projects were completed, the professionals were starting to care about different data, aspects of the patients' lives and organizations that they did not connect to so far.

The new entanglements did lead to new frictions and questions about representation. There were many discussions about how to present the machine learning models to their colleagues and about introducing them to the mental healthcare practice. These discussions showed how deeply the distribution of epistemic responsibility is reconfigured when engaging with machine learning. Often, the professionals felt that the details of their entanglements with the machine learning models were lost within "simple" percentages and graphs. How could they explain all the nuances and intricacies of their models? And if they could not, how can they be responsible?

During the last training day, professionals D and F led a plenary discussion about presenting their findings. Both professionals struggled; they had nearly completed the course and their organizational projects, but felt that they needed to know more and do more before they were comfortable incorporating their models into healthcare practice:

*"Professional F: "The danger with this machine learning is that we, again, are going to do many things automatically, and then I miss again knowledge of what I am doing. You have to know the program so well!" Professional D: "I run a script and I get an outcome, but I often do not have an idea about what I am comparing. I find that complicated. I have no idea that I am doing what is described in theory." [...] D: "This is also scary, that you can do it so wrong!" (Observation day 4)*

The decisions that the professionals make, were tied to certain situations, outcomes and, as the above example illustrates, the professionals find it "scary" to present a black box to their organizations, which hides the responsibilities, interdependencies and vulnerabilities. They hoped that the course would bring them better, easier knowledge but that turned out not to be the case: by engaging in new relations, they experienced anew the political and ethical aspects of caring about and for data, processes and outcomes.

In these discussions, often a particular configuration of responsibility was put forth. Teacher Z continually suggested further engagement rather than detachment in order to make these decisions. He recommends engaging with the domain experts and investigating with

them what it is they care about, in order to use those insights to shape machine learning and towards better knowledge.

Responsibility in the sense of better explainable models, means being accountable for the reason for one's knowledge. During the course, the overall tendency was to be responsible through engagement with other human actors, rather than by assigning more responsibility to machine learning based on (minimally) better performance. Rather than only look at the criterion of prediction in evaluating a model, a more explainable, more accountable model was preferred over a model that made slightly better predictions but was more opaque in its functioning and less open to engagement.

These many instances show how being responsible in a machine learning trajectory is all about finding a careful balance, building new networks and including others. There was a constant need to interest others in the predictions of machine learning, and to account and to form new relations between colleagues and the machine learning approach.

## Discussion

The insights we gained from our fieldwork contribute to an alternative understanding and investigation of responsibility that contrasts with individualized conceptualizations of responsibility in philosophy of technology. Our work highlights the importance of situated ethics and knowledge, and the crucial role that technologies have in knowledge production. In contrast to many popular accounts of machine learning, we find that machine learning does not reveal radical new knowledge, but it does include some mutual adjustments in knowledge, technology and ethics, the formation of new entanglements and involves intense commitments to taking and delegating responsibility.

In particular, we highlight how a particular kind of introduction to machine learning can support professionals in their understanding of responsibility as situated. The teacher continually stressed the importance of including the particularities of a case and continually learning with others. Furthermore, the moments when machine learning feels like a 'breach' in the normal script of responsibility reveals 'responsibility-as-usual' to the participants. This could be an important learning moment that is cultivated, a privileged moment to stay with the trouble (Haraway, 2016), rather than an awkward phase that one has to go through to move on to machine learning wisdom. Such frictions signal differences in ways of taking responsibility across epistemic cultures, as epistemic virtues are negotiated (Stevens et al., 2020). This opens up important opportunities for reflection in machine learning trajectories.

Our view is that attention to how humans and non-humans are engaged in caring both reveal patterns about kinds of responsibility and point us to the opportunities, the spaces for avoiding neglect and irresponsibility. Our approach allows the importance of friction come into focus. This shows the added value of our analysis: rather than call for control and clarification, and deplore that mediation creates even more irresponsible black boxes, a care approach shows how responsibility-in-the-making can emphasize ethical reflection and action. Our project also raises questions with regards to the mechanisms that reduce this friction. As Edwards and colleagues have noted (Edwards et al., 2011), it is important to pay attention to how we smooth things and to the way algorithms can be valued in different ways —without necessarily reaching closure (Lee and Helgesson, 2020).

The promises of prediction and of upscaling are part of the drivers of the adoption of machine learning in mental healthcare. The standardization of data and analyses and the application of generalized rules would pave the way for a faster out roll of machine learning. Yet, these come at the cost of the entanglements that make machine learning work in what our participants considered a responsible way. We need to remain aware and critical of the possible scaling effects, of the consequences of the concentration of power, and the negative effects of non-accountable knowledge-producing techniques. By focusing on responsibility as the result of entanglements and stressing the importance of ongoing engagements and iterative learning processes, we shed light on these issues.

## Footnotes

<sup>1</sup> Epistemic responsibilities are the responsibilities that actor(s) take in the process of knowing. Together, these responsibilities shape responsible knowledge practices.

<sup>2</sup> See. Pitt (2014) and Hanson (2014) for more information about methodological individualism.

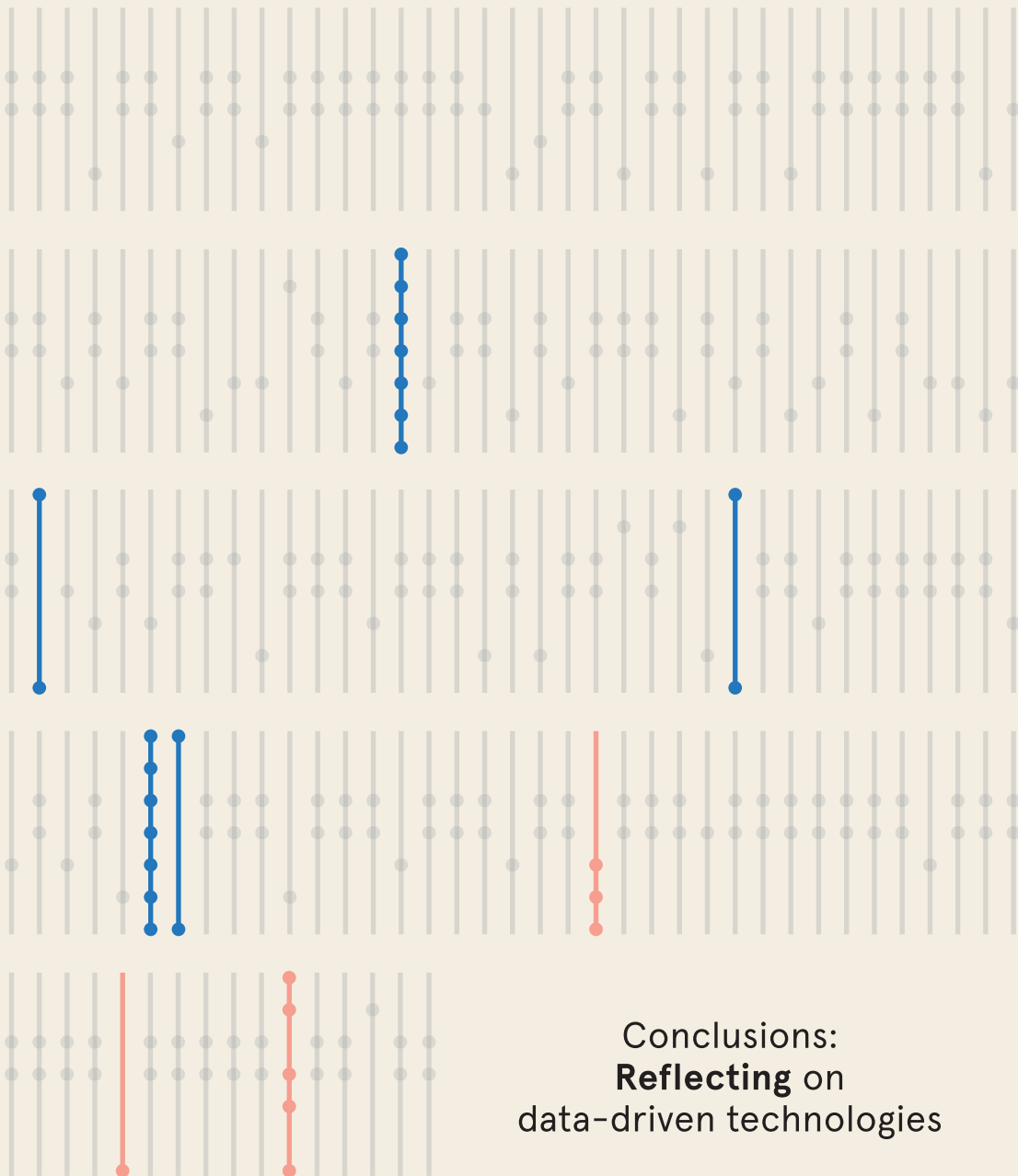
<sup>3</sup> Also, care is not necessarily positive or benevolent. “Bad” technologies, such as bombs, polluting cars and nerve gas also need care to keep functioning. This is also the same for data; data that is not cared for does not get used and does not endure (Leonelli, 2014; Beaulieu, forthcoming).

<sup>4</sup> An ethical waiver was obtained for the study. We asked continued informed consent from all the participants and incorporated member checks in various phases of the study.

<sup>5</sup> We conducted nine semi-structured interviews with participating healthcare professionals to explore topics, such as the self-study and the organizational project, in more depth. In addition, all participants filled in a questionnaire before and after the course, to provide more insight into the general expectations of the training, and into the experience of all participants. All responses and fieldnotes were expanded upon, transcribed, analyzed and discussed as soon as possible by MS and AB.



# Chapter 7



Conclusions:  
**Reflecting on**  
data-driven technologies





## Introduction

*"All epistemology begins in fear – fear that the world is too labyrinthine to be threaded by reason; fear that the senses are too feeble and the intellect too frail; fear that memory fades, even between adjacent steps of a mathematical demonstration; fear that authority and convention blind; fear that God may keep secrets and demons deceive" (Daston and Galison, 2007: 374).*

Many data-driven technologies envisaged through hopes and fears are based on assumptions about the relation between technology and epistemology. Some people dream that data-driven technology will fundamentally change epistemic practices in healthcare. They hope that the new data can be used to make better predictions, smarter decisions and more personalized interventions. Other people, however, fear that data-driven technology will lead to incomplete, inscrutable and misguided decision-making with possibly detrimental effects.

The aim of this dissertation is to critically investigate how data-driven technologies reconfigure responsible knowledge practices in healthcare. Over the years, epistemic practices have been developed in healthcare, guided by particular norms and virtues as to what is deemed responsible evidence, information and decision-making. Data-driven technologies promote the disruption of the existing epistemic cultures in the healthcare sector. This means that healthcare actors must rethink and reposition their epistemic practices as they need to relate to these new technologies and decide what data, information and knowledge is responsible to use in healthcare.

This dissertation takes an emic perspective in terms of what is deemed a responsible knowledge practice. It views epistemology as something enacted in practice rather than as an external category or a set of abstract criteria. It means that epistemic practices cannot be separated from actions and doings. In healthcare, people do not explicitly discuss "epistemology", but they value particular information, decide what evidence to use and negotiate on methodology. Such practices give insight into the epistemic cultures of healthcare fields, in defining which norms and values play a role, in deciding what is *responsible* knowledge and in terms of the actions that follow from this.

In the introduction I argued that this emic perspective is a necessary addition to current discussions about the permissibility of data-driven healthcare. First, adopting an emic perspective allows me to move beyond the artificial distinction between ethics and epistemology as it accounts for the interrelatedness of both (Jasanoff, 2004; Simon, 2015). Second, because it explicitly considers the embedded normativities and affectivities that

are already playing a role in healthcare practices (cf. Zuiderent-Jerak, 2007). Current ethical literature does excellent work in (1) agenda-setting for particular ethical and epistemological dilemmas, (2) theorization by focusing on particular normative and epistemic trade-offs and (3) problematization by highlighting how data-driven technologies misalign with our current ethical concepts and principles. Yet, these approaches risk introducing overarching normative elements from the outside that do not do justice to the normativities in the studied practices themselves.

The main research question I addressed in this dissertation is: *How do various actors reconfigure responsible knowledge practices in data-driven healthcare?*

This main question is divided into the following sub-questions:

1. *How do actors in data-driven healthcare envision responsible knowledge practices?*
2. *How do actors in data-driven healthcare construct responsible knowledge practices?*
3. *How are roles and responsibilities with regards to knowledge practices in data-driven healthcare reconfigured?*

Like a scavenger, I looked around and picked up diverse traces to gain an understanding of how responsible knowledge practices were reshaped based on data-driven technologies. I needed to move flexibly from site to site, using diverse methods to gather and analyze my data. All data were gathered between December 2016 and July 2019 and in total 164 interviews were combined with more than 250 hours of observations and extensive document analyses.

In the previous chapters, I assembled accounts of various epistemic practices in healthcare which have introduced data-driven technologies either by envisioning or by day-to-day construction. I traced the envisioning of data-driven healthcare back to writing in scientific journals from the healthcare domain (Chapter 2) and the accounts of diverse experts involved in Big Data initiatives stimulated by the European Union (Chapters 3 and 4). I tracked down the construction of data-driven technologies to a pioneering hospital-based initiative in the Netherlands (Chapter 5) and a machine learning training program for medical professionals that stimulated the normalization of the technologies (Chapter 6).

Here I will answer the research questions, based on the findings of the previous chapters, and reflect on the answers in terms of existing scholarly work. Then I will discuss the implications this study has for theory, practice, and methodology and follow this with suggestions for future research. The chapter ends with some personal reflection and concluding remarks.

## Envisioning responsible knowledge practices

For the first sub-question, I set out to study how actors envision responsible knowledge practices in data-driven healthcare. The reason for including the envisioning of data-driven technologies is that the expectations surrounding data-driven technologies are filled with many epistemic claims. I used the concepts of discourse (Chapter 2) and ethical framing (Chapter 3) and the metaphor of dreams (Chapter 4) to explore various aspects of envisioning that surround data-driven technologies.

Chapter 2 describes three dominant positive discourses that seem to equate data-driven technologies with responsible knowledge practices. Two-thirds of the editorials on Big Data in scientific journals in the healthcare domain between 2012 and 2016 are positive about the shift toward Big Data. These discourses seem to reinforce one another in the idea that data-driven technologies are highly reliable; they frequently highlight the many benefits for knowledge production. The most extreme version (the modernist discourse) posits that current epistemic practices in the healthcare sector are even unsustainable and outdated, suggesting that it is irresponsible *not* to use the new technological opportunities.

Scholarly editorials pay less attention to the scientific and critical-interpretive discourses than to the positive discourses. These more critical discourses argue that Big Data could also lead to irresponsible knowledge practices as it too easily discards scientific methodologies, or because it ignores the context of data, leading to biased, wrong conclusions.

The strong presence of positive discourses discussed in Chapter 2 can explain the drive toward data-driven healthcare. Many scholars emphasize the performative effect of the expectations surrounding emerging technology. They say that such expectations do not merely describe reality but profoundly influence the role and function of technology in our society (e.g. Konrad et al., 2017; Jasanoff and Kim, 2015; Van Lente, 2012). This led me to conclude that the strength of the positive discourses between 2012 and 2016 and their assumptions about improved knowledge production offer important explanations for the many data-driven initiatives that have started since then (e.g. KPMG, 2020; Jiang et al., 2017; Rieder, 2018).

Chapter 3 addresses how the promise steers us toward particular responsible knowledge practices, explicitly with the notion of ethical framing. Actors, sometimes without being aware of it, distribute roles and responsibilities in their depictions of how ethics should deal with data-driven technology (cf. Cool, 2019; Hoeyer et al., 2017). The chapter thus broadens the focus on how actors envision responsible knowledge practices by taking into account the role of ethical framing. We distinguish three ethical frames: (1) Framing

ethics as a ‘balancing act’ mobilizes policymakers and other stakeholders as rational actors who should engage in utilitarian decision-making by finding the optimal balance between contrasting ethical principles. (2) Framing ethics as a ‘technical fix’ mobilizes technicians, policymakers and policymakers to develop clear guidelines and instrumental solutions to ‘solve’ ethical concerns. In contrast, (3) framing ethics as a ‘collective thought process’ mobilizes affected stakeholders and even citizens more broadly to become involved in deliberative processes and steer the agenda toward forms of deliberation. This chapter shows that the specific ways of framing ethics steer the promises and expectations toward particular responsible knowledge practices. *What* responsibility in this context means and by whom it should be enacted thus depends on the ethical frame.

By focusing on the experts working in data-driven initiatives, Chapter 4 provides deeper insights into why positive discourses seem to prevail. Other scholars have described how individuals flexibly connect their images for technological futures to their own past and present experience (e.g. Konrad et al., 2017; Van Lente, 2012). We wanted to gain insight into such epistemic-affective sensemaking processes by applying the metaphor of dreams to look at the experts’ motivations, values and considerations. The chapter addresses how experts link the grand promise of data-driven technology to their own experience of the generalized, scattered, slow and uncontrollable nature of health information. The experts also seem to connect data-driven technologies explicitly to desires and needs that are arguably fundamentally human: being seen, feeling connected, feeling secure, and feeling protected against harm. This suggests that the way data-driven technologies are envisaged offers much flexibility for experts to link them to their own needs and wishes for improving epistemic practices.

Chapter 4 also highlights how experts working on data-driven initiatives nuance the promises surrounding data-driven technology. They make trade-offs in the way they envision data-driven healthcare, depending on their motives, values and considerations. The chapter clearly shows that actors make such trade-offs. For example, in describing their dream of tailoring healthcare information to individual patient’s needs, the experts realized that it must be made easier to disclose personal information, thereby portraying a particular idea of a responsible knowledge practice needed to achieve their dream.

To sum up, strong positive discourses surrounding data-driven technologies see opportunities for improving epistemic practices in the healthcare sector and promote the notion that using technology is responsible. Healthcare actors make sense of such promises by flexibly linking the promises to their own experiences and desires for improving knowledge practices. At the same time, the promises are far from neutral as they hide all sorts of trade-offs and ethical frames ascribe particular roles and responsibilities to actors.

## Constructing responsible knowledge practices

To answer my second sub-question, I aimed to investigate ethnographically how actors construct responsible knowledge practices in data-driven initiatives in the healthcare sector, inspired by the positive discourses that surround data-driven technologies. Chapters 5 and 6 report on my studies of such initiatives in the mental healthcare field. I identified three ways in which actors create responsible knowledge practices.

First, actors construct responsible knowledge practices by *creating separation* between data-driven technologies and healthcare-as-usual. Chapter 5 reports on a data-driven initiative by the psychiatry department of a large Dutch hospital. This initiative created various forms of separation between the technologies and healthcare-as-usual. It founded a new “innovation” team, recruiting data scientists from outside the (mental) healthcare sector. Framed as a “pioneering” initiative that experiments with the use of these technologies, it was set apart from the regular, day-to-day care provided in the department. Such separation is also visible in Chapter 6, where healthcare professionals learn the basics of machine learning and work on a machine learning project for their own organization. The machine-learning course they were on repeatedly framed the initiatives as “test cases”.

Creating separation from “fields-out-there” is a phenomenon extensively addressed in the STS literature on experimentation and laboratory studies (e.g. Knorr Cetina, 1992; 1995). This body of work describes how researchers start an experiment to study how, for instance, a particular technology might work and to showcase its possibilities (e.g. Lezaun et al., 2017). The idea is that secluding – in this case – of machine learning experimentation from healthcare-as-usual helps to establish what is viewed as the “truth” about the workings of the technology in relative separation from all the irregularities, exceptions and noise of routine healthcare practice (cf. Knorr Cetina, 1995; Ivanova, 2020).

The healthcare professionals and data scientists understood that they were separating data-driven technologies from healthcare-as-usual. In Chapters 5 and 6, they kept talking about the necessary conditions for introducing data-driven technologies to healthcare practices. In Chapter 5, the psychiatrists negotiated with data scientists on what kind of “certainty” was needed before medical decisions could be based on the prediction models (that were based on machine learning techniques). One conclusion was that the models should be tested extensively against care-as-usual. In Chapter 6, discussions centered on when and how to present the outcomes of the prediction models to other therapists uninvolved in the development process. The healthcare professionals who developed the models found it important to be transparent to their peers about the choices they had made for the model and to not steer healthcare practice too little or too much in any particular direction. These

examples serve to illustrate that for these healthcare professionals, introducing data-driven technologies to healthcare routines is only responsible when the technology has been tested in full and are accompanied by “the right amount” of contextual information.

Second, the actors construct responsible knowledge practices by *adapting* the data-driven technologies to specific healthcare settings. The actors understood that introducing data-driven technologies to healthcare practice was more than a simple matter of implementation. They realized that they must go through a process of translating data-driven technologies to the particularities of their settings (cf. Janssen, 2016; Latour, 1987). Chapters 5 and 6 highlight that knowledge practices of the healthcare field led the data-driven initiatives. Psychiatrists (Chapter 5), and “domain experts” (psychiatrists and therapists in Chapter 6) played an important role in corroborating the work of data scientists and the outcomes of the prediction models.

Chapter 5 gave an example of how psychiatrists and therapists adapted data-driven technologies to their practices. The first case was typical in revealing the different approaches to shaping data-driven technologies. The data scientists and a psychiatrist tried to negotiate on the sort of data to include for an antidepressant prediction model. The data scientists took a flexible, inclusive approach to selecting data and questioned the need for cut-off points and threshold levels that the psychiatrist regularly used. In contrast, the psychiatrist was strict on data inclusion –not all data were of “good” quality – and upheld the importance of maintaining the threshold levels. The chapter shows how the actors negotiated their differences in situ. They worked through the various types of data and possible threshold levels and made decisions for them all. In these discussions, the psychiatrist upheld many of his current norms, which meant that the data scientists had to adapt their methods to the epistemic practices of the healthcare field.

Third, the actors construct responsible knowledge practices by *specifying* the use of data-driven technologies. Healthcare actors ascribe data-driven technologies to specific issues and questions in which their role is deemed responsible, without letting the technologies “overflow” to areas where their use is uncertain.

The actors in Chapters 5 and 6 come to understand data-driven technologies as playing two roles. According to the first role, data-driven technologies can help psychiatrists or therapists to develop adequate diagnoses and treatments. They can use the technologies to predict potential side effects of medication (Chapter 5) or to estimate which patients will likely drop out of addiction treatment (Chapter 6). The idea is that this kind of insight can be a timely signal for psychiatrists and therapists, for intervention or to adjust treatment to the individual patient’s needs. These settings that are currently known as highly uncertain,

meaning that there is little information or evidence to base decisions on. But the professionals maintain a leading role in treatment trajectories and must be able to ignore any signals. According to the second role, data-driven technologies can enhance the efficiency of healthcare organizations with regard to “safe” topics. The technologies can be used to schedule staff, organize patient beds and predict treatment costs and time (Chapter 6). These sorts of topics are deemed safe, as they do not directly interfere with patient care and are currently surrounded by much uncertainty.

To conclude, in contrast to the views put forward in positive discourses, the actual construction of data-driven healthcare is a complex, slow process. Healthcare actors assess and study the possibilities of data-driven technologies in relative separation from care-as-usual. They only consider introducing the technologies to the healthcare field once the innovation has been properly tested and adapted to the epistemic practices of their particular field, and its use is restricted to areas of healthcare that are considered “responsible”. In this way, healthcare professionals tame data-driven technologies to fit their epistemic practices (cf. Smits, 2002).

## Reconfiguring roles and responsibilities

The final sub-question asked how roles and responsibilities are reconfigured with regards to knowledge practices in data-driven healthcare. Scholars have described how knowledge is produced in highly entangled, complex networks combining humans and non-human agents, such as databases, technologies, infrastructures and organizations (Callon and Law, 1997; Latour, 1987; Simon, 2015). I conclude that knowledge-production networks become more diverse and complex with data-driven technologies. As the network changes, actors may rethink their roles and responsibilities and adapt their epistemic practices (cf. Smits, 2002).

First, new actors are introduced to the knowledge-producing networks in healthcare and join the negotiations on knowledge. The empirical chapters describe how data scientists (and data engineers, information scientists, privacy officers) who used to be (partly) on the outside of care are now exerting a substantial influence on knowledge production in the healthcare sector. Data scientists, for example, suggest what data could be gathered and bring in their own understanding of epistemic virtues (Chapter 5). These actions affect the way knowledge is produced in the network as they lead to in situ ‘trading’ (Chapter 5) and the internalization of new ways of working (Chapter 6).

Second, new technologies are added to the networks. Chapter 6 describes how data-driven technologies help shape the knowledge-producing network. We gave the example of “Lasso”, a machine learning algorithm “wrapper” that can be used in R (a programming language and computing environment) to support analyses. It pre-sorts the analysis by making specific selections. This example shows that technology also operates as an actor when it comes to knowledge production.

Responding to changes in the network, human actors such as psychiatrists (Chapter 5) began to rethink their own roles and responsibilities. As they encountered the new actors, they were asked to explain parts of their epistemic culture that they had not thought about for a long time. Sometimes they began questioning their own epistemic practices. In Chapter 5, the psychiatrists faced other ways of determining “certainty” than they were familiar with. These sorts of encounters made the healthcare actors aware of their own epistemic practices and opened opportunities for change.

Similarly, data scientists started to rethink their own roles and responsibilities and adapted their practices to healthcare fields. In Chapter 5, the data scientist learned that healthcare data were more challenging to analyze than they had initially envisaged. Increasingly they understood that healthcare fields could not be as easily transformed as, for instance, the call center of a commercial travel company. Realizing that they needed specific understanding of healthcare, they studied medical guidelines to learning what was necessary for responsible knowledge practices in the healthcare field and adapting the data-driven technologies. These findings show that the epistemic cultures in healthcare do not exist in a vacuum but have flexible borders and that epistemic virtues are open to changing interpretations (Daston and Galison, 2007; Galison, 1997; Smits, 2002).

At the same time, power relations are clearly visible in these networks (cf. Puig de la Bellacasa, 2007). The chapters show that the data-driven technologies are received in a sector with strong ideas about norms and values, roles and responsibilities. The healthcare actors were not simply impressed by the potential of data-driven technologies and kept their practices as leading. Especially the psychiatrists and therapists (Chapters 5 and 6) had powerful roles in deciding on data, methodologies and outcomes. This meant that many of the existing roles and responsibilities in healthcare fields with regards to knowledge production were reinforced and maintained.

To sum up, I observed some reconfiguration of roles and responsibilities in knowledge practices in healthcare. The most important finding is that the knowledge-producing network becomes more complex and existing relations are reshaped with the introduction of new human and technological actors. As a result of the changes in networks, healthcare actors



and data scientists reflect on their own roles and responsibilities. However, the power to decide on epistemic practices in healthcare remained in the hands of certain healthcare actors.

## Overall conclusion

The answers to the sub-questions helped me answer the main question of this study: *how do various actors reconstruct responsible knowledge practices in data-driven healthcare?*

Contrary to the modernist vision of radical revolution with the sudden arrival of new information and knowledge practices, achieving data-driven healthcare is proving to be a slower and more collaborative process. Envisaged through shared hopes and personal dreams, healthcare actors invite data-driven technologies and data scientists to experiment with data-driven healthcare in their field as they assume it has potentially beneficial effects for their epistemic practices.

When looking at the practices, the chapters show that data-driven technologies are not being introduced in an empty playing field. They arrive in settings in which all sorts of norms and values continue to play strong roles. Much discussion and negotiation about responsible knowledge production takes place, following a pragmatist approach as actors do not choose data, virtues and responsibility a priori, but ad hoc, according to the circumstances (cf. Smits, 2002).

Throughout the chapters, three points characterize the current way of constructing responsible knowledge practices in data-driven healthcare: First, responsible knowledge practices are reconstructed by *keeping a human-centered focus*. The development of data-driven healthcare is often surrounded by stories of uncontrollable automated decision models that take over knowledge production. This dissertation shows that in the data-driven initiatives in healthcare that I studied, human understanding and control are perceived as necessary conditions. Healthcare actors want to understand every stage of knowledge production based on data-driven technologies and are very careful in handing over their capacities of knowledge production to technologies.

Second, actors reconstruct responsible knowledge practices with a degree of *reluctance to change epistemic practices too abruptly*. Knowledge-producing networks become more diverse and complex in data-driven healthcare initiatives. The human actors that I studied in Chapters 5 and 6 were confronted with their own epistemic virtues and assumptions about knowledge. This made them reflect on their roles and responsibilities in produc-

ing knowledge, a process that opened up opportunities for change. However, given their cautious approach to taking unnecessary risks, these healthcare actors remained most powerful in producing healthcare knowledge and kept their epistemic practices in the lead of initiatives.

Third, responsible knowledge practices are reconstructed by *specifying the role of data-driven technologies in particular settings*. Healthcare actors consider some areas of healthcare more suitable for data-driven innovation than others. This dissertation shows that data-driven technologies are used for specific issues that currently are highly uncertain or have manageable and overseeable consequences. In this situation, the role of the technologies is deemed responsible.

To sum up, reconstructing data-driven healthcare is a slow process in which actors actively shape responsible knowledge practices based on three mechanisms: by maintaining a human-centered focus, being cautious not to change epistemic practices too abruptly and specifying the role of data-driven technologies in particular settings.

## Theoretical implications

This dissertation adds in many ways to existing scholarly work and theories. For example, it contributes to our understanding of the gradual uptake of data-driven technologies in healthcare, by offering ethnographic insight into the day-to-day, mundane work in data-driven healthcare initiatives in healthcare. According to many scholars, such work is needed to interrogate the extreme hopes and fears that surround the technologies (e.g. Ames, 2018; Crawford et al., 2014; Kennedy, 2018; Wehrens et al., 2020). However, I would like to dedicate this part of the dissertation to describing how my *approach* to studying data-driven technologies contributes to existing scholarly work on data-driven innovation.

## Ethics-in-practice

This dissertation adds to existing ethics literature by approaching ethics as an empirical phenomenon. In the introduction, I wrote that much of the ethical literature on data-driven technologies focuses on agenda-setting, theorization, or problematization (e.g. Floridi and Sanders, 2004; Grote and Berens, 2020; Mittelstadt and Floridi, 2016; Mittelstadt et al., 2016; Mittelstadt, 2019; Morley et al., 2020). This is valuable work but also separates ethics from the concrete practices in which ethical decisions are made and ethical work is done. This separation often continues in practice by appointing “ethical committees” and defining “rules and regulations” that portray ethics as a phase one has to go through or

something that can be reached by ticking all the boxes on a “responsible machine learning” checklist (e.g. The Institute for Ethical AI & Machine Learning, 2020).

The value of approaching ethics as an empirical phenomenon is that it more precisely shows what norms and values are at stake in situ. This dissertation shows that technologies are not introduced to an “empty” healthcare field, but a field that has rich epistemological traditions and in which many epistemic norms and values play a role. Data-driven technologies add another layer to these discussions, negotiations began on the responsible use of the technologies and conclusions are drawn about the “right” thing to do in a particular setting. These discussions offer valuable insights into the normativities deemed important in the field and the choices made can be used for ethical reflection.

This perspective also implies that “doing ethics” is not just reserved for ethical scholars and committees, but that ethical decisions are made and responsible knowledge practices are created by everyone working with data-driven technologies (e.g. Ames, 2018). Ethical scholars and experts have a role in articulating and explicating what ethical decisions are made and showing the assumptions and tensions involved. In sum, therefore, my dissertation highlights the need for a descriptive, emic approach to ethics that can serve as a useful addition to prescriptive work that focuses on defining rules and ethical principles upfront.

## **Epistemology-in-practice**

This dissertation adds in a similar way to discussions about epistemology. My dissertation focuses on epistemology-in-practice when looking at data-driven technologies. In the introduction I argued that tending to responsible knowledge practices adds to how the scholarly literature currently addresses epistemology. Ethical literature treats epistemology as an abstract category, for example, to theorize on the inconclusive, inscrutable and misguided nature of evidence that can result from data-driven technologies (Lipworth et al., 2017; Mittelstadt et al., 2016; Morley et al., 2020). Or to discuss how deploying data-driven technologies might shift evidentiary norms of medical diagnosis (Grote and Berens, 2020).

The epistemology-in-practice perspective moves beyond such theorizations as it gives empirical insight into a variety of epistemic concerns and normativities at stake in data-driven initiatives. It offers a nuanced view that picks up on small changes in epistemic practices and differences in epistemic cultures.

For example, many ethics scholars discuss the opacity or inscrutable nature of data-driven technologies. Opacity is deemed problematic as medical doctors cannot assess the reliability of prediction models, which can lead to inappropriate or biased conclusions (e.g. Grote

and Berens, 2020; Morley et al., 2020; Wachter, 2015). This dissertation nuances these discussions as it shows that current data-driven technologies can be opaque in different ways and for different people. Chapter 6 shows, for instance, that technologies can take over part of the selection process of variables. Such opacity is not considered problematic by all involved, because such choices are currently already supported by many statistical tools.

The epistemology-in-practice perspective also highlights how politics is entangled with knowledge production. Such concerns are often neglected by current literature on ethical data science. To be clear, this literature is concerned with ethical themes such as bias and accountability but neglects the broader entanglement of politics and knowledge production (cf. Moats and Seaver, 2019). Adopting this epistemology-in-practice perspective provides new insights and concerns. For example, this dissertation concludes that healthcare actors (in Chapters 5 and 6, psychiatrists and therapists) lead the decision-making process on epistemic practices in healthcare. Such findings produce new questions, such as: is it right that healthcare actors play this role in determining what responsible knowledge practices are? Who else should be involved?

## Theoretical conceptualizations

This dissertation puts forward diverse concepts that can be used to study responsible knowledge practices in data-driven healthcare: ethical frames, epistemic dreams, epistemic cultures (Knorr Cetina, 1981;1999), epistemic virtues (Daston and Galison, 2007), trading zones (Galison, 1997), epistemic responsibility (Simon, 2015).

Two of these concepts were developed within this dissertation. The concept of *ethical framing* allows us to study how ethical arguments become tied to the promises of data-driven healthcare technologies. The concept brings together STS literature on promissory futures, ethics literature on data-driven technologies, and policy literature on framing as discursive work. The concept adds especially to ethics literature by revealing how, in promises, actors discursively conduct ethics work, by assigning roles and responsibilities.

I introduced the metaphor of *dreams* to study the personal sensemaking processes and underlying drivers, values and considerations for actors involved in data-driven initiatives. It offers another perspective on knowledge practices than found in STS, CAS and CDS literature that stresses the social dynamics of expectations. This concept adds especially to CAS and CDS literature, emphasizing the actors' normativities and affectivities in data-driven initiatives.

These six concepts can be used to study the construction of responsible knowledge practices in data-driven healthcare. Collectively, these concepts provide a rich and valuable vocabulary to explore how data-driven technologies are slowly becoming part of healthcare practices.

## Practical implications

This dissertation describes how responsible knowledge practices are created in data-driven healthcare. The studied practices show that there is no simple way of deriving “responsible knowledge practices”. What is deemed responsible is always determined in particular settings and in interaction with many actors. I have taken inspiration from the work of Puig de la Bellacasa (2017) on *care* and Huijer (2018) on *expansive thinking* (in Dutch: *uitgebreid denken*) to stimulate the development of responsible knowledge practices in data-driven healthcare. These are my practical implications:

### There is time for reflection

The first practical implication concerns the perceived speed of development. Many hyped technologies and buzzwords have been used to signal the trend toward intensified data use in healthcare. They promise radical transformation and speedy disruption of existing knowledge practices, suggesting that healthcare actors and organizations need to hurry or otherwise they will miss out on the developments and become obsolete (cf. De Wilde, 2000).

However, this dissertation argues that we tend to overestimate the pace of change in healthcare. I focused on data-driven healthcare practices instead of individual hyped technologies and their capabilities. This perspective reveals that many of the ideas behind data-driven healthcare are less innovative than often thought. Chapter 4 illustrates that many data-driven dreams can be associated with age-old frustrations in the healthcare sector and fundamental human needs. In addition, shaping data-driven healthcare takes time as it requires much negotiation, collaboration and reflection between different actors.

The implication for healthcare regulators, managers and professionals is that there is time, and this time is necessary to reflect on how data-driven technologies can support responsible knowledge practices. Deciding on responsible knowledge practices is not something for one person or one group of actors to do. This dissertation shows that knowledge is created in networks in which many actors rely upon each other. When only one person or group of actors decides what is “responsible” it is likely to lead to disagreement, misunderstanding and missed opportunities for epistemic reflection.

I recommend arranging broad conversations with medical doctors, nurses, patients, researchers, data scientists, IT personnel, administrators and all other actors involved in the knowledge-producing networks. Conversation topics could include, for example, the data-driven dreams of actors (Chapter 4) and the epistemic uncertainties that actors face (Chapter 5). Together, actors could look for similarities, find solutions for frustrations and problems, and decide how their knowledge practices can best be supported (possibly by data).

Knowledge-producing networks contain both human and non-human actors. Therefore, it is just as important to get to know the expertise of the technologies involved so that the strengths of all actors can be used. It can be helpful to imagine how technologies care for responsible knowledge practices (see also the example of the worms in Chapter 6). We can learn to involve their perspective, attend to their needs and ensure their contribution to responsible knowledge practices too (cf. Huijer, 2018; Puig de la Bellacasa, 2017).

### **Make data-driven initiatives more resilient**

This dissertation gives insight into how responsible data-driven healthcare is currently shaped by actors in the field. However, there is much uncertainty about the influence of data-driven technologies on future knowledge practices. Responding to uncertainty, we can choose to introduce rules and regulations in order to determine responsibilities in advance.

At the same time, rules and regulations *alone* are not desirable for two reasons. First, rules and regulations can start to lead a life on their own and lose connection with the field that they aim to protect. Examples include the debate on regulatory pressure in the Dutch healthcare sector. Healthcare professionals protest against the great deal of registration work that they must do to comply with regulations as it hinders their capacity to provide good care (Felder, 2020). Second, simply because it is impossible to regulate everything. Healthcare practices often change faster than the regulators can manage to create new regulations.

I argue that it is important to make data-driven healthcare initiatives more resilient (cf. Bal et al., 2015; Mesman, 2011). This dissertation shows that actors in the healthcare field have internalized a strong set of norms and values and have formed strong ideas about responsible knowledge. Instead of appointing and fixing responsibilities in rules, it is important to explore the “texture” (Mesman, 2011: 72) of responsible data-driven healthcare. Achieving new responsible knowledge practices is not only based on error-reducing activities but also on understanding and learning from the responsibilities already present in the field.

I would advise healthcare regulators and managers to be careful with introducing too many rules and regulations. Instead, I would recommend proceeding by identifying and sharing best practices and encouraging the development of platforms where people can share experiences. For example, by organizing focus groups and data meetings.

Social scientists and philosophers could help to stimulate resilient data practices and data-driven initiatives by encouraging conversations, bringing in new perspectives and articulating opposing views. Empirical research can help to reveal underlying dreams, ways of framing and expose important principles that might have been backgrounded to the everyday workings of data-driven healthcare practices.

### **Ensure continuity in data-driven initiatives**

The third implication of this dissertation concerns the experimental character of many data-driven initiatives. This study found that as many data-driven initiatives begun in (relative) separation of the healthcare field, extensive negotiation and collaboration are required to make the technologies fit particular healthcare practices. Based on this study, I would recommend ensuring continuity in data-driven initiatives to facilitate such negotiations. This means two things:

First, healthcare regulators, managers and professionals should be careful not to separate data-driven initiatives from concrete healthcare practices in which the technology will eventually be used. The study shows that data-driven technologies are always adapted to the particular epistemic practices of a healthcare setting and vice versa. Expecting data-driven technology to be a simple matter of “copy-paste” from one healthcare setting to another ignores the complex interactions important to developing responsible data-driven practices.

Second, if data-driven technologies are used in healthcare, they require ongoing attention, maintenance and repair (cf. Puig de la Bellacasa, 2017). This work is necessary to ensure that the technologies continue to fit the intended practices. This means that it is important to invest in continued, in-house expertise, for example, by hiring data scientists or training medical professionals to work with these technologies.

### **Develop appropriate rules to safeguard data-driven initiatives**

Lastly, it is important to develop appropriate rules to safeguard the settings in which data-driven initiatives take place. This means introducing and evaluating rules that not only support the collective experimental development process but also take into account the needs of patients and professionals (Bal et al., 2015).

Such rules can be very functional in offering opposition to power blocks. This dissertation shows that negotiation and collaboration on responsible knowledge practices are crucial in developing data-driven healthcare. It is important to firmly counter concentration of power that can lead to all sorts of asymmetries in data-driven healthcare (Sharon, 2016; 2018). For example, large technological corporations that may become such powerful actors they limit the freedom of individual actors or organizations in the field. In this case, negotiations can no longer proceed, and resilient practices are under pressure. Appropriate rules are essential to creating the conditions that enable responsible knowledge practices.

## Methodological implications

Methodologically, this dissertation shows that one can study fleeting, nebulous technologies while they are being developed and (start to) reshape locally embedded knowledge practices. The CAS and CDS literature often portray data-driven technologies as powerful yet inscrutable entities that somehow govern, shape and control our lives. I noticed that the technologies are often packed with complex terminology, statistical theories and a lot of hype. However, by reiterating the complex nature of this phenomenon, we also end up mystifying a phenomenon that we seek to clarify (Ziewitz, 2017). This dissertation shows that it is possible to study something that is widely thought to be inscrutable.

For years we have been doing “multi-sited”, “distributed” and “networked” ethnography (Burrell, 2009; Hine, 2007; Howard, 2002), following objects around that do not stay in one place. This offers a more in-depth understanding of phenomena than a single case can provide. We often combine multi-sited approaches with multiple or mixed methodologies, such as diverse quantitative and qualitative methods.

As a variant of these methodologies, the “scavenger” metaphor has been used in CAS (Seaver, 2017). It highlights a flexible approach focused on assembling different traces, allowing the analyst to move flexibly across disciplinary boundaries and combine and experiment with methodologies to explore the role of data-driven technologies. What becomes visible is an object of study that is less subjected to the hypes of the day and not caught up in the boundaries of particular disciplines and the methods that they use.

The scavenger metaphor advocates using and experimenting with different ethnomethods. In my eyes, such experimentations are a valuable *addition* to the methodologies that we normally use. Developing the illustrations for Chapter 2 and the text mining tool I used in Chapter 4 required me to work together with people from several disciplines. These people presented data to me in different lights. Thus, reflection across domains is helpful not only



for the actors in our fields, but *also* for us researchers to deepen reflection and enhance “our epistemic practices”.

I find it important to emphasize that more flexibility in selecting cases and developing methods does not mean that anything goes. It is and remains important to work together with other researchers (remember; scavengers operate in troops or packs), have member checks in place that help to put the traces you find into context, and reflect on your role as a researcher in your setting, for example, by organizing interdisciplinary cooperation.

## Strengths and limitations

In this dissertation, I aimed to navigate between all the hopes and fears that surround data-driven technologies in popular discourses. This stance allowed me to explore the nuanced subtle shifts that data-driven technologies bring to the healthcare field. However, this is also a normative position. It highlights current practices and takes those as a starting point for further development. At the same time, one could argue, that this positions backgrounds more structural negative consequences, such as systematic bias and discrimination practices. We need to remain aware and critical of such effects. At the same, I believe that this dissertation contributes by showing the many ways in which data-driven technologies are steered in healthcare, not only by rules and regulations, but also by cultures, norms and values.

I chose to focus on the overarching data-driven technologies in general, instead of focusing on particular technologies. The risk of this approach is that it seems to equalize all the diverse technologies, while neglecting differences (for example, between supervised and unsupervised machine learning approaches). This broader focus is warranted, I believe, as in the introduction, I noticed that the technological labels followed each other in rapid succession but continued to promise similar things. Making this distinction is helpful in studying overarching trends and strategies.

Another issue to reflect on is that both my ethnographic case studies were conducted in the mental healthcare field, which could arguably lead to conclusions that are specific to this field only. The rationale behind zooming in on this field in particular have to do with historical characteristics. As the mental healthcare field has a history of resistance to quantification and as the success factors for treatment are largely unknown, my assumption was that the discussions surrounding data-driven technologies (both in terms of perceived risks and in terms of promise to reduce epistemic uncertainty) would be amplified. Next to these ethnographic cases, the interviews, document analysis and review of editorials that

form the basis of chapters 2, 3 and 4 focused on other fields in medicine as well. I believe that the conclusions reported in this thesis therefore carry wider saliency than the mental healthcare field.

A final point to reflect upon relates to the various concepts that are developed in Chapters 2, 3 and 4. The notions of 'dreams', 'framing' and 'discourses' all focus primarily on the envisioning dimension of data-driven technologies. One might argue that these chapters create a somewhat artificial distinction, leading to different overviews that perhaps could be captured under one term. However, such critique would neglect the subtle differences of focus. I believe that it is the combination of concepts that has most value, as the different overviews highlight slightly different things, therefore allowing a richer understanding of the dimension of 'envisioning'. While the notion of 'discourses' highlights how different assumptions and metaphors form a coherent whole, the concept of 'dreams' allowed me to capture underlying drivers that motivate individual people to invest time and energy in data-driven initiatives. The notion of ethical framing adds relevant insights into how ethical themes are envisaged and with what consequences for roles and responsibilities of other actors.

## Future research

The outcomes of this study provide various opportunities for future research. This study could be expanded in several ways; it points to areas that need further attention. But allow me to use this opportunity to present the outline of an agenda for studying data-driven healthcare.

I recommend that social scientists and philosophers take a pro-active, ongoing role in studying technological developments *while* they are being developed. In the field, ethical and social reflection is too often viewed as a "phase" that initiatives or people must go through, like "getting approval" from an ethical committee, for example. Responsibility becomes something achievable, and irresponsibility can be solved instrumentally or technically. A consequence of this way of approaching the issues is that many uncertainties and issues that accompany technological innovation turn into procedures that must be solved instrumentally (Wehrens et al., 2019). As researchers, we enable this kind of thinking by presenting "simple" rules and guidelines for the ethical and responsible use of technologies.

Instead, social scientists and philosophers should take a more continuous role in helping actors with these initiatives to develop "responsibility-in-practice". They should not play

the referee (cf. Huijjer, 2018), or provide a clear framework for responsible data practices. Instead, we should stimulate continuous reflection, provide order in overflowing normativities (Zuiderent-Jerak, 2007) and stimulate discussion. Here, empirical research helps us see what is going on, it gives professionals a voice and reveals their doubts and caution.

Currently, there is much uncertainty about how to deal with data-driven technologies. It is impossible to oversee all the issues beforehand. While this often leads to a precautionary response, I would like to plea, following Verbeek, for a focus on technology ‘accompaniment’: continuously following technological developments, studying the people, data, methods, cultures, norms and values and providing feedback to help “guide” the technologies to find a good place in our society (Verbeek, 2014; Smits, 2002). This is how we can carefully help translate those data technologies developed in relative separation from healthcare (see answer to sub-question 2) to practices in the field.

I would suggest always doing this in the specificity of the field. I do not recommend increasing the differences between healthcare fields and data-driven technologies by starting up labs or creating extra experimental settings. A recent example lab is the creation of ELSA (ethical legal societal aspects) labs, in which different partners come together to experiment on and test the acceptance of technologies. In my opinion, these sorts of initiatives only increase the separation from the day-to-day work in which the technologies eventually will take place. Often the idea is that what works in the lab will likely work in other settings. Instead, I would argue for continual experimentation and development of technologies for *particular* settings.

## Personal reflection

I am writing the closing words of this dissertation during the Covid-19 pandemic. In this time of uncertainty, people fear for their health, feel disconnected and are unsure about the future. It has led to people expressing their own ideas about the way forward, critiquing and questioning scientific studies and protesting openly against governmental policies. In other times and other places, we might have valued the sudden interest in science and politics. Now, it just seems to be driving people apart just when we need each other the most.

In this context, I am presenting a dissertation on responsible knowledge practices that describes different epistemic cultures, each with their own norms and virtues, existing side by side. I am explicitly not taking sides and argue that epistemic differences should be seen as a starting point for rethinking and evaluating our own epistemic culture.

This might be considered a dangerous combination. Therefore, I would like to emphasize that I believe that the findings and conclusions of this study can be used to arrange for *more* responsible knowledge practices in healthcare. We should move beyond polarized discussions and encourage reflection on how knowledge is obtained in healthcare. It is my position that epistemic practices can always be improved by (critical) reflection; correcting the wrong or unintended consequences of practices is always valuable.

Important, in my eyes, especially in these times, in moving forward is to connect and present the various actors and positions. I chose to study data-driven technologies across different disciplines, actors and organizational boundaries. And soon this dissertation, which started out as a reflection on epistemic practices, also became about interdisciplinary cooperation and making difficult topics a matter of collective reflection.

This fits with my own interdisciplinary background, ranging from medicine to health law, from philosophy to economy and from social science to healthcare governance. I have always been averse to disciplinary niches and never wanted to commit myself too strictly to any field. This is reflected in my research. During my PhD, I was trained in STS – an interdisciplinary research field – but on the way I could not resist including the philosophy of technology, CAS and CDS insights (and absorb a thing or two about machine learning).

During this PhD, I also learned that bringing different perspectives together is not always easy. Occasionally, I met resistance against this research project. I frustrated some medical professionals who did not understand the distinctions I drew between the epistemic practices of healthcare and those of data science. Data-driven technologies were just a lot of hype in their eyes. They had already been doing lots of large-scale predictive analyses in the medical field “for ages and doing them better.” Others kindly told me that I should learn more about statistical learning before I could say anything about data-driven technologies; some of their arguments reminded me of the discipline wars between qualitative and quantitative research.

This resistance, these questions, influenced the research project. It showed me that questions about epistemology are very *fundamental* for people (as the Daston and Galison quote at the beginning of this chapter also illustrates). Even asking questions about knowledge, information and evidence was easily experienced as intimidating and interpreted as criticism, which was not my intention. It led me to develop less-threatening approaches to discuss epistemology. I hope this dissertation shows that it is possible to reflect meaningfully on complex and obscure themes such as “responsible knowledge practices.”

I take this as my responsibility in future research: to keep looking for ways to bring people, perspectives and understandings together; to always try to understand the other first before I form my own ideas and, wherever possible, encourage interdisciplinary cooperation. Because, if anything, Covid-19 has shown us all the importance of connecting to others.

## Concluding remarks

This is the end of this dissertation. A reader looking for easy answers to the construction of responsible knowledge practices with data-driven technology might be disappointed. A part of me wishes that the issues we face in healthcare could be solved by in ten simple steps or by applying an algorithm. Unfortunately, there are no easy answers to complex problems and this dissertation cannot provide easy answers.

However, this dissertation provides ample insight into the current construction of responsible knowledge practices in healthcare. It highlights the social and technological nature of our ways of knowing and shows the practices, norms and values that play a role. It also describes how welcoming interaction with new actors and methodologies leads to reflection on existing epistemic practices, and shows that our data dreams can inspire us to create *more* responsible knowledge practices in healthcare every day.



# References





- Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A and Havel J (2013) Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine* 11(2): 47–58.
- Ames MG (2018) Deconstructing the algorithmic sublime. *Big Data & Society* 5(1): 1–4.
- Ananny M (2016) Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Science, Technology, & Human Values* 41(1): 93–117.
- Anderson C (2008) The end of theory; the data deluge makes the scientific method obsolete. In: *Wired*. Available at: [www.wired.com/2008/06/pb-theory/](http://www.wired.com/2008/06/pb-theory/) (accessed 5 September 2017).
- Andrejevic M (2020) Shareable and un-sharable. *Big Data & Society* 7(1): 1–4.
- Aradau C and Blanke T (2015) The (Big) Data-security assemblage: Knowledge and critique. *Big Data & Society* 2(2): 1–12.
- Asaro PM (2019) AI ethics in predictive policing: From models of threat to an ethics of care. *IEEE Technology and Society Magazine* 38(2): 40–53.
- Baird D and Cohen MS (1999) Why Trade? *Perspectives on Science* 7(2): 231–254.
- Bal R (2017) Evidence-based policy as reflexive practice. What can we learn from evidence-based medicine? *Journal of Health Services Research & Policy* 22(2): 113–119.
- Bal R, Stoopendaal AMV and van de Bovenkamp H (2015) Veerkracht en veiligheid: hoe kan regulering van de zorg daaraan bijdragen. *Nederlands Tijdschrift van de Geneeskunde* 159(1): 1–3.
- Balogh EP, Miller BT and Ball JR (2015) *Improving Diagnosis in Health Care*. Washington: National Academies Press.
- Barad K (1996) Meeting the universe halfway: Realism and social constructivism without contradiction. In: Hankinson N and Nelson J (eds) *Feminism, Science, and the Philosophy of Science*. Dordrecht: Kluwer, pp. 161–194.
- Barnes TJ and Wilson MW (2014) Big Data, social physics, and spatial analysis: The early years. *Big Data & Society* 1(1): 1–14.
- Barrett D (2020) *Pandemic Dreams*. Cambridge: Oneiroi Press.
- Baru C (2019) Data Science Needs a Clinical Program. Available at: [www.chaitanbaru.blogspot.com/2019/02](http://www.chaitanbaru.blogspot.com/2019/02) (accessed March 2020).
- Basak SC, Vracko M and Bhattacharjee AK (2015) Big Data and New Drug Discovery: Tackling ‘Big Data’ for Virtual Screening of Large Compound Databases. *Current Computer-Aided Drug Design* 11(3): 197–201.
- Bauman Z and Lyon D (2013) *Liquid surveillance: A conversation*. Hoboken: John Wiley & Sons.
- Beaulieu A (2001) Voxels in the brain: neuroscience, informatics and changing notions of objectivity. *Social Studies of Science* 31(5): 635–680.
- Beaulieu A and de Rijcke S (2014) Networked Neuroscience: Brain Scans and Visual Knowing at the Intersection of Atlases and Databases. In: Coopmans C and Woolgar S (eds) *Representation in scientific practice revisited*. Cambridge: The MIT Press, pp. 366.
- Beaulieu A and Leonelli S (Forthcoming) *A Critical Introduction to Data and Society*. London: SAGE.
- Beer D (2016) How should we do the history of Big Data? *Big Data & Society* 3(1): 1–10.
- Beer T and Grote P (2019) On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics* 46(3): 205–211.
- Belgrave D, Simpson A and Custovic A (2014) Challenges in Interpreting Wheeze Phenotypes: The Clinical Implications of Statistical Learning Techniques. *American Journal of Respiratory and Critical Care Medicine* 189(2): 121–123.
- Bensaude Vincent B (2014) The politics of buzzwords at the interface of technoscience, market and society: The case of ‘public

- engagement in science'. *Public Understanding of Science* 23(3): 238-253.
- Berghout M (2020) *Medical leadership: transforming professional practice and identity*. Phd Thesis, Erasmus Universiteit Rotterdam, NL.
- Berghout MA, Oldenhof L, van der Scheer WK and Hilders CG (2020) From context to contexting: professional identity un/doing in a medical leadership development programme. *Sociology of Health & Illness* 42(2): 359-378.
- Berwick DM (2016) Era 3 for Medicine and Health Care. *JAMA* 315(13): 1329-1330.
- BigMedilytics (2018) Infographics. Available at: [www.eur.nl/eshpm/media/2019-01-workpackage53infographicsfinalversions](http://www.eur.nl/eshpm/media/2019-01-workpackage53infographicsfinalversions) (accessed 28 October 2020).
- Bigmedilytics (2020) Improving healthcare in Europe through Big Data. Available at: [www.bigmedilytics.eu](http://www.bigmedilytics.eu) (accessed 30 October 2020).
- Bijker WE (2006) Why and how technology matters. In: Goodin RE and Tilly C (eds) *Oxford handbook of contextual political analysis*. Oxford: Oxford University Press, pp.1-26.
- Borup M, Brown N, Konrad K and Van Lente H (2006) The sociology of expectations in science and technology. *Technology Analysis & Strategic Management* 18(3-4): 285-298.
- boyd d and Crawford K (2012) Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662-679.
- Brown ML (2016) Can't you just pull the data? The limitations of using of the electronic medical record for research. *Pediatric Anesthesia* 26(11): 1034-1035.
- Brown N and Michael M (2003) A sociology of expectations: Retrospecting prospects and prospecting retrospects. *Technology Analysis & Strategic Management* 15(1): 3-18.
- Burkeley K (2019) Dreaming is Imaginative Play in Sleep: A theory of the function of dreams. *Dreaming* 29(1): 1-21.
- Burrell J (2009) The Field Site as a Network: A Strategy for Locating Ethnographic research. *Field Methods* 21(2): 181-199.
- Burrows R and Savage M (2014) After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data & Society* 1(1): 1-6.
- Busch L (2017) Looking in the Wrong (La)place? The Promise and Perils of becoming Big Data. *Science, Technology, & Human Values* 42(4): 657-678.
- Callon M and Law J (1997) After the individual in society: lesson on collectivity from science, technology and society. *Canadian Journal of Sociology* 22(2): 165-182.
- Cases M, Furlong LI, Albanell J, Altman RB, Bellazzi R, Boyer S, Brand A, Brookes AJ, Brunak S, Clark TW, Gea J, Ghazal P, Graf N, Guigó R, Klein TE, López-Bigas N, Maojo V, Mons B, Musen M, Oliveira JL, Rowe A, Ruch P, Shabo A, Shortliffe EH, Valencia A, van der Lei J, Mayer MA and Sanz F (2013) Improving data and knowledge management to better integrate health care and research. *Journal of Internal Medicine* 274(4): 321-328.
- Chalmers MK and Edwards PN (2017) Producing "one vast index": Google Book Search as an algorithmic system. *Big Data & Society* 4(2): 1-16.
- Collins H and Evans R (2008) *Rethinking Expertise*. Chicago: The University of Chicago Press.
- Collins H, Evans R and Gorman M (2007) Trading zones and interactional expertise. *Studies in History and Philosophy of Science* 38(4): 657-666.
- Cool A (2019) Impossible, unknowable, accountable: Dramas and dilemmas of data law. *Social Studies of Science* 49(4): 503-530.
- Crawford K, Miltner K and Gray ML (2014) Critiquing Big Data: Politics, Ethics,

- Epistemology. *International Journal of Communication* 8(1): 1663-1672.
- Cruz TM (2020) Perils of data-driven equity: Safety-net care and Big Data's elusive grasp on health inequality. *Big Data & Society* 7(1): 1-14.
- Dalton C and Thatcher J (2014) What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'Big Data'. In: Space and Society Open Site. Available at: [www.societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/](http://www.societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/) (accessed 6 April 2017).
- Daston L and Galison PL (2007) *Objectivity*. New York: Zone Books.
- De Wilde R (2000) *De voorspellers: een kritiek op de toekomstindustrie*. Amsterdam: De Balie.
- Demiaux V and Si Abdallah Y (2017) *how can humans keep the upper hand? The ethical matters raised by algorithms and artificial intelligence*. Report. December 2017. Paris: Commission Nationale de l'Informatique et des Libertés
- DeVan A (2016) The 7 V's of Big Data. In: Impact Radius. Available at: [www.impactradius.com/blog/7-vs-big-data/](http://www.impactradius.com/blog/7-vs-big-data/) (accessed 11 September 2017).
- Deville J, Guggenheim M and Hrdličková Z (2016) Same, same but different: Provoking relations, assembling the comparator. In: Sahin O (eds) *Practising comparison: logics, relations, collaborations*. Manchester: Mattering Press, pp. 99-129.
- Dixon-Wood M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, HSU R, Katbamna S, Olsen R, Smith L, Riley R and Sutton AJ (2006) Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology* 6(35): 1-13.
- Dwyer D, Falkai P and Koutsouleris N (2018) Machine learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology* 14(1): 91-118.
- Ebeling MF (2016) *Healthcare and Big Data*. Philadelphia: Palgrave Macmillan.
- EC (2014) European Commission and data industry launch €2,5 billion partnership to master Big Data. Available at: [www.ec.europa.eu/commission/presscorner/detail/en/IP\\_14\\_1129](http://www.ec.europa.eu/commission/presscorner/detail/en/IP_14_1129) (accessed 30 October 2020).
- Edwards PN, Mayernik MS, Batcheller AL, Bowker GC and Borgman CL (2011) Science Friction: Data, Metadata, and Collaboration. *Social Studies of Science* 41(5): 667-690.
- Egbert S (2018) About discursive storylines and techno-fixes: the political framing of the implementation of predictive policing in Germany. *European Journal for Security Research* 3(2): 95-114.
- Epstein S (1996) *Impure science: aids, activism and the politics of knowledge*. Chicago: The University of California Press.
- Felder M and Meerding WM (2017) *Een toekomst voor Evidence-Based Medicine? Achtergrondstudie bij het advies "Zonder context geen bewijs. Over de illusie van evidence-based practice in de zorg."* Report. June 2017. The Hague: Raad voor Volksgezondheid en Samenleving.
- Felt U (2015) Keeping technologies out: Sociotechnical imaginaries and the formation of Austria's technopolitical identity. In: Jasanoff S and Kim SH (eds) *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power* Chicago: The University of Chicago Press, pp. 103-125.
- Floridi L and Sanders JW (2004) On the morality of artificial agents. *Minds and Machines* 14(3): 349-379.
- Freeman JV and Saxon L (2015) Remote Monitoring and Outcomes in Pacemaker and Defibrillator Patients. *Journal of the American College of Cardiology* 65(24): 2611-2613.

- Galison PL (1995) Introduction: The Context of Disunity. In: Galison PL and Stump DJ (eds) *The Disunity of Science: Boundaries, Contexts and Power*. Palo Alto: Stanford University Press, pp. 1-36.
- Galison PL (1997) *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.
- Galison PL and Stump DJ (1995) *The Disunity of Science: Boundaries, Contexts and Power*. Palo Alto: Stanford University Press.
- Gallagher F (2013) Is Big Data the next Big Brother? In: Wired. Available at: [www.wired.com/insights/2013/03/is-big-data-the-next-big-brother/](http://www.wired.com/insights/2013/03/is-big-data-the-next-big-brother/) (accessed 25 October 2020).
- Garrison LP (2013) Universal Health Coverage - Big Thinking versus Big Data. *Value in Health* 16(1): S1-S3.
- Geels FW and Verhees B (2011) Cultural legitimacy and framing struggles in innovation journeys: a cultural-performative perspective and a case study of Dutch nuclear energy (1945-1986). *Technological Forecasting and Social Change* 78(6): 910-930.
- Ghani KR, Zheng K, Wei JT and Friedman CP (2014) Harnessing Big Data for Health Care and Research: Are Urologists Ready? *European Urology* 66(6): 975-977.
- Gillespie T (2020) Content moderation, AI, and the question of scale. *Big Data & Society* 7(2): 1-5.
- Giraud EH (2019) *What Comes after Entanglement: Activism, Anthropocentrism, and an Ethics of Exclusion*. Durham: Duke University Press.
- Goffman E (1974) *Frame analysis: An essay on the organization of experience*. Cambridge: Harvard University Press.
- Gomella LG (2016) Big Data Equals Big Challenges for Prostate Cancer. *The Canadian Journal of Urology* 23(4): 8329.
- Greenhalgh T, Howick J and Maskrey N (2014) Evidence based medicine: a movement in crisis? *BMJ* 348: g3725.
- Grote T and Berens P (2020) On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics* 46(3): 205-211.
- Grunwald A (2004) The normative basis of (health) technology assessment and the role of ethical expertise. *Poiesis & Praxis* 2(2-3): 175-193.
- Haan MM, van Gurp JLP, Naber SM and Groenewoud AS (2018) Impact of moral case deliberation in healthcare settings: a literature review. *BMC Medical Ethics* 19(85): 2-15.
- Hanson FA (2014) Which Came First, the Doer or the Deed? In: Kroes P and Verbeek PPCC (eds) *The Moral Status of Technical Artefacts*. Dordrecht: Springer Science & Business Media, pp. 55-73.
- Haraway DJ (2016) *Staying with the trouble: Making kin in the Chthulucene*. Durham: Duke University Press.
- Harmen O and Galison PL (2008) Epistemic Virtues and Leibnizian Dreams: On the Shifting Boundaries between Science, Humanities and Faith. *The European Legacy* 13(5): 551-575.
- Harris A, Kelly SE and Wyatt S (2014) Autobiologies on YouTube: narratives of direct-to-consumer genetic testing. *New Genetics and Society* 33(1): 60-78.
- Hay SI, George DB, Moyes CL and Brownstein JS (2013) Big Data Opportunities for Global Infectious Disease Surveillance. *PLoS Medicine* 10(4): 1-4.
- Hayes P, van de Poel I and Steen M (2020) Algorithms and values in justice and security. *AI & Society* 35: 533-555.
- Heeney C (2017) An "ethical moment" in data sharing. *Science, Technology, & Human Values* 42(1): 3-28.
- Hine C (2007) Multi-sited Ethnography as a Middle Range Methodology for Contemporary STS. *Science, Technology, & Human Values* 36(6): 652-671.
- Hoeyer K (2019) Data as promise: Reconfiguring Danish public health through personal-

- ized medicine. *Social Studies of Science* 49(4): 531-555.
- Hoeyer K, Tupasela A and Rasmussen MB (2017) Ethics policies and ethics work in cross-national genetic research and data sharing: flows, nonflows, and overflows. *Science, Technology, & Human Values* 42(3): 381-404.
- Househ M, Aldosari B, Alanazi A, Kushiruk AW and Borycki EM (2017) Big Data, Big Problems: A Healthcare Perspective. *Informatics Empowers Healthcare Transformation* 238: 36-39.
- Howard PN (2002) Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods. *New Media & Society* 4(4): 550-574.
- Huijer M (2018) *Denkende verbeelding: Publieksfilosofie in de 21<sup>ste</sup> eeuw*. Amsterdam: Boom Uitgevers.
- Huxley A (1931) *Brave New World*. London: Vintage.
- Hyysalo S, Pollock N and Williams RA (2019) Method matters in the social study of technology: Investigating the biographies of artifacts and practices. *Science & Technology Studies* 32(3): 2-25.
- Introna LD (2014) Towards a post-human intra-actional account of sociomaterial agency (and morality). In: Kroes P and Verbeek PPCC (eds) *The Moral Status of Technical Artefacts*. Dordrecht: Springer Science & Business Media, pp. 31-53.
- Ivanova D (2020) *Care/Place: Unsettling place in healthcare*. PhD Thesis, Erasmus Universiteit Rotterdam, NL.
- Janssen M (2016) *Situated Novelty: A study on healthcare innovation and its governance*. PhD Thesis, Erasmus Universiteit Rotterdam, NL.
- Jasanoff S (2004) *States of knowledge: the co-production of science and the social order*. London: Routledge.
- Jasanoff S and Kim SH (2009) Containing the atom: Sociotechnical imaginaries and nuclear power in the United States and South Korea. *Minerva* 47(2): 119-146.
- Jasanoff S and Kim SH (2015) *Dreamscapes of modernity: sociotechnical imaginaries and the fabrication of power*. Chicago: University of Chicago Press.
- Jerak-Zuiderent S (2015) Keeping open by re-imagining laughter and fear. *The Sociological Review* 63(4): 897-921.
- Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H and Wang Y (2017) Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology* 2(4): 230-243.
- Kellogg KC, Orlikowski WJ and Yates J (2006) Life in the Trading Zone: Structuring Coordination across Boundaries in Post-bureaucratic Organizations. *Organization Science* 17(1): 22-44.
- Kennedy H (2018) Living with Data: Aligning Data Studies and Data Activism Through a Focus on Everyday Experiences of Datafication. *Krisis: Journal for Contemporary Philosophy* 1(1): 18-30.
- Khoury MJ and Ioannidis JPA (2014) Big Data meets public health. *Science* 346(6213): 1054-1055.
- Kitchin R (2014) Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 1-12.
- Kitchin R and Lauriault TP (2014) Towards critical data studies: Charting and unpacking data assemblages and their work. In: Eckert J, Shears A, and Thatcher J (eds) *Geoweb and Big Data*. Lincoln: University of Nebraska Press, pp. 3-23.
- Kitchin R and McArdle G (2016) What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society* 3(1): 1-10.
- Klonoff D (2013) Twelve Modern Digital Technologies That Are Transforming Decision Making for Diabetes and All Areas of Health Care. *Journal of Diabetes Science and Technology* 7(2): 291-296.

- Knorr Cetina K (1981) *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press.
- Knorr Cetina K (1995) Laboratory Studies: The Cultural Approach to the Study of Science. In: Jasanoff S, Markle GE, Petersen JC and Pinch T (eds) *Handbook of Science and Technology Studies*. London: SAGE, pp. 140-166.
- Knorr Cetina K (1999) *Epistemic Cultures: How the sciences make knowledge*. Cambridge: Harvard University Press.
- Koed Madsen A (2018) Data in the smart city: How incongruent frames challenge the transition from ideal to practice. *Big Data & Society* 5(2): 1-3.
- Konrad K, van Lente H, Groves C and Selin C (2017) Performing and Governing the Future in Science and Technology. In: Felt U, Fouché R, Miller CA and Smith-Doerr L (eds) *The Handbook of Science and Technology Studies fourth edition*. Cambridge: The MIT Press, pp. 465-494.
- Kottyan LC, Weirauch MT and Rothenberg ME (2015) Making it big in allergy. *Journal of Allergy and Clinical Immunology* 135(1): 43-45.
- KPMG (2020) *Inventarisatie AI-toepassingen in gezondheid en zorg in Nederland: Onderzoek naar de stand van zaken in 2020*. Report. September 2020. Amstelveen: KPMG.
- Krakoff LR and Phillips RA (2016) Blood Pressure Variability. *Journal of the American College of Cardiology* 68(13): 1387-1388.
- Kremer JAM, Noordegraaf-Eelens LHJ and Ottes L (2017) *Zonder context geen bewijs*. Report. June 2017. The Hague: Raad voor Volksgezondheid en Samenleving.
- Kroes P and Verbeek PPCC (2014) Introduction. In: Kroes P and Verbeek PPCC (eds) *The Moral Status of Technical Artefacts*. Dordrecht: Springer Science & Business Media, pp. 1-9.
- Kruse CS, Goswamy R, Raval YJ and Marawi S (2016) Challenges and opportunities of Big Data in health care: a systematic review. *JMIR medical informatics* 4(4): 1-11.
- Kuhn T (1962) *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Lakoff G and Johnson M (2011) *Metaphors we live by: with a new afterword*. Sixth edition. Chicago: University of Chicago Press.
- Latour B (1987) *Science in Action*. Cambridge: Harvard University Press.
- Latour B (2004) Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. *Critical Inquiry* 30(2): 225-248.
- Latour B and Woolgar S (1986) *Laboratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press.
- Lee F and Helgesson CF (2019) Styles of Valuation: Algorithms and Agency in High Throughput Bioscience. *Science Technology, and Human Values* 45(4): 659-685.
- Lee F and Larsen LB (2019) How should we theorize algorithms?: Five ideal types in analyzing algorithmic normativities. *Big Data & Society* 6(2): 1-6.
- Lee F, Bier J, Christensen J, Engelmann L, Helgesson CF and Williams R (2019) Algorithms as folding: Reframing the analytical focus. *Big Data & Society* 6(2): 1-12.
- Leem JG (2016) Big Data and pain. *The Korean Journal of Pain* 29(4): 215-216.
- Leonelli S (2014) What difference does quantity make? On the epistemology of Big Data in biology. *Big Data & Society* 1(1): 1-11.
- Levenda AM, Richter J, Miller T and Fisher E (2019) Regional sociotechnical imaginaries and the governance of energy innovations. *Futures* 109(1): 181-191.
- Lezaun J, Marres N and Tironi M (2017) Experiments in Participation. In: Felt U, Fouché R, Miller CA and Smith-Doerr L (eds) *The Handbook of Science and Technology*

- Studies fourth edition*. Cambridge: The MIT Press, pp. 195-222.
- Loke YK and Derry S (2003) Does anybody read 'evidence-based' articles? *BMC Medical Research Methodology* 3(1): 1-6.
- MacRae CA (2012) Pattern recognition: combining informatics and genetics to re-evaluate conduction disease. *Heart* 98(17): 1263-1264.
- Madsen AK and Munk AK (2019) Experiments with a data-public: Moving digital methods into critical proximity with political practice. *Big Data & Society* 6(1): 1-14.
- Maiers C (2017) Data Talk: Metaphors of an Epistemological Landscape. In: *Annual Meeting of the Society for Social Studies of Science* Boston, USA, August 30 - September 2, 2017
- Marcum JA (2017) Clinical Decision-making, Gender Bias, Virtue Epistemology, and Quality Healthcare. *Topoi* 36(3): 501-508.
- Marcus GE (1995) Ethnography in/of the world system: The emergence of multi-sited ethnography. *Annual Review of Anthropology* 24(1): 95-117.
- Marr B (2014) Big data: The 5 vs everyone must know. In: LinkedIn. Available at: [www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know](http://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know) (accessed 25 October 2020).
- Mayer-Schönberger V and Cukier K (2014) *Big data: a revolution that will transform how we live, work, and think*. Boston: Mariner Books, Houghton Mifflin Harcourt.
- McAfee A and Brynjolfsson E (2012) Big Data: The Management Revolution. *Harvard Business Review*. 90(10): 60-68.
- McNutt TR, Moore KL and Quon H (2016) Needs and Challenges for Big Data in Radiation Oncology. *International Journal of Radiation Oncology\*Biophysics\*Physics* 95(3): 909-915.
- Mesman J (2011) Resources of strength: an exnovation of hidden competences to preserve patient safety. In: Rowley E and Waring J (eds) *A socio-cultural perspective on patient safety*. Farnham: Ashgate Publishing, pp. 71-94.
- Metcalfe J and Crawford K (2016) Where are human subjects in Big Data research? The emerging ethics divide. *Big Data & Society* 3(1): 1-14.
- Micheli M, Ponti M, Craglia M and Berti Suman A (2020) Emerging models of data governance in the age of datafication. *Big Data & Society* 7(2): 1-15.
- Miller FA, Ahern C, Smith CA and Harvey EA (2006) Understanding the new human genetics: A review of scientific editorials. *Social Science & Medicine* 62(10): 2373-2385.
- Mittelstadt B (2019) The Ethics of Biomedical 'Big Data' Analytics. *Philosophy & Technology* 32(1): 17-21.
- Mittelstadt B, Allo P, Taddeo M, Wachter S and Floridi L (2016) The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(2): 1-21.
- Mittelstadt BD and Floridi L (2016) The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics* 22(2): 303-341.
- Moats D and Seaver N (2019) "You Social Scientists Love Mind Games": Experimenting in the "divide" between data science and critical algorithm studies. *Big Data & Society* 6(1): 1-11.
- Mol A (2002) *The body multiple: Ontology in medical practice*. Durham: Duke University Press.
- Mol A (2006) Proving or improving: on health care research as a form of self-reflection. *Qualitative Health Research* 16(3): 405-414.
- Mor Y (2014) Big Data and law enforcement: was 'minority report' right? In: Wired. Available at: [www.wired.com/insights/2014/03/big-data-law-enforcement-minority-report-right/](http://www.wired.com/insights/2014/03/big-data-law-enforcement-minority-report-right/) (accessed 25 October 2020).
- Morley J, Machado CCV, Burr C, Cows J, Joshi I, Taddeo M and Floridi L (2020) The ethics of AI in health care: A mapping review.



- Social Science & Medicine* 260(113172): 1-14.
- Mortelmans D (2007) *Handboek kwalitatieve onderzoeksmethoden*. Den Haag: Acco.
- Murdoch TB and Detsky AS (2013) The inevitable application of Big Data to health care. *JAMA* 309(13): 1351-1352.
- Najarian K, Ward KR and Shirani S (2013) Bio-medical Signal and Image Processing for Clinical Decision Support Systems. *Computational and Mathematical Methods in Medicine* 2013(1): 1-2.
- Narula J (2013) Are We Up to Speed? *JACC: Cardiovascular Imaging* 6(11): 1222-1224.
- Nicolini D, Powell J, Conville P and Martinez-Solano L (2008) Managing knowledge in the healthcare sector. A review. *International Journal of Management Reviews* 10(3): 245-263.
- Nissenbaum H (1998) *Values in Computer System Design: Bias and Autonomy: Ethics and Information Technology*. Delhi: New Academic Press.
- Oldenhof L (2015) *The Multiple Middle: Managing in Healthcare* Phd Thesis, Erasmus Universiteit Rotterdam, NL.
- Oldenhof L and Wehrens R (2018) Who is 'in' and who is 'out'? Participation of older persons in health research and the interplay between capital, habitus and field. *Critical Public Health* 28(3): 281-293.
- Oldenhof L, Stoopendaal A and Putters K (2016) Professional talk: How middle managers frame care workers as professionals. *Health Care Analysis* 24(1): 47-70.
- O'Neil C (2016) *Weapons of Math Destruction*. New York: Broadway Books.
- Onukwugha E (2016) Big Data and Its Role in Health Economics and Outcomes Research: A Collection of Perspectives on Data Sources, Measurement, and Analysis. *PharmacoEconomics* 34(2): 91-93.
- Orlikowski WJ and Gash DC (1994) Technological frames: making sense of information technology in organizations. *ACM Transactions on Information Systems (TOIS)* 12(2): 174-207.
- Orwell G (1949) 1984. London: Secker & Warburg.
- PacMed (2019) Projects. Available at: [www.pac-med.ai/nl/projects/](http://www.pac-med.ai/nl/projects/) (accessed September 2019).
- Patrick K (2016) Harnessing Big Data for health. *Canadian Medical Association Journal* 188(8): 555.
- Paul H (2011) Performing history: how historical scholarship is shaped by epistemic virtues. *History and Theory* 50(1): 1-19.
- Penkler M, Felder K and Felt U (2019) Challenging Diversity: Steering Effects of Buzzwords in Projectified Health Care. *Science, Technology, & Human Values* 45(1): 138-163.
- Pigliucci M (2017) The Virtuous Skeptic. *Skeptical Inquirer: The Magazine For Science and Reason* 41(2): 1.
- Pinch TJ and Bijker WE (1984) The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science* 14(3): 399-441.
- Pine KH and Bossen C (2020) Good organizational reasons for better medical records: The data work of clinical documentation integrity specialists. *Big Data & Society* 7(2): 1-13.
- Pitt JC (2015) "Guns Don't Kill People, People Kill"; Values in and/or Around Technologies. In: Kroes P and Verbeek PPCC (eds) *The Moral Status of Technical Artefacts*. Dordrecht: Springer Science & Business Media, pp. 89-101.
- Pollock N and Williams R (2010) The business of expectations: How promissory organizations shape technology and innovation. *Social Studies of Science* 40(4): 525-548.
- Pope C, Halford S, Tinati R and Weal M (2014) What's the big fuss about 'Big Data'? *Journal of Health Services Research & Policy* 19(2): 67-68.
- Potters L, Ford E, Evans S, Pawlicki T and Mutic S (2016) A Systems Approach Using Big



- Data to Improve Safety and Quality in Radiation Oncology. *International Journal of Radiation Oncology\*Biological\*Physics* 95(3): 885–889.
- Prainsack B (2018) The “we” in the “me” solidarity and health care in the era of personalized medicine. *Science, Technology, & Human Values* 43(1): 21–44.
- Prainsack B (2019) Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society* 6(1): 1–15.
- Priault N, Weinel M and Wrigley A (2016) Rethinking moral expertise. *Health Care Analysis* 24(4): 393–406.
- Puig de la Bellacasa M (2017) *Matters of Care*. Minneapolis: University of Minnesota Press.
- Puschmann C and Burgess J (2014) Metaphors of Big Data. *International Journal of Communication* 8(1): 1690–1709.
- Raghupathi W and Raghupathi V (2014) Big Data analytics in healthcare: promise and potential. *Health Information Science and Systems* 2(1): 1–10.
- Rathenau Institute (2018) Science in Figures: Infrastructure of knowledge. Available at: [www.rathenau.nl/en/science-figures/policy-and-structure/infrastructure-knowledge](http://www.rathenau.nl/en/science-figures/policy-and-structure/infrastructure-knowledge) (accessed March 2020).
- Rein M and Schön DA (1993) Reframing policy discourse. In: Fischer F and Forester J (eds) *The argumentative turn in policy analysis and planning*. Durham: Duke University Press, pp. 145–166.
- Restifo NP (2013) A ‘Big Data’ View of the Tumor ‘Immunome’. *Immunity* 39(4): 631–632.
- Rexhepi P (2016) Liberal luxury: Decentering Snowden, surveillance and privilege. *Big Data & Society* 3(2): 1–3.
- Ribes D (2019) STS, Meet Data Science, Once Again. *Science, Technology, & Human Values* 44(3): 514–539.
- Rieder G (2018) Tracing Big Data Imaginaries through Public Policy: The Case of the European Commission. In: Rudinow Sætnan A, Schneider I and Green N (eds) *The Politics and Policies of Big Data: Big Data, Big Brother?*. London: Routledge, pp. 89–109.
- Rieder G and Simon J (2016) Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society* 3(1): 1–6.
- Risoud M, Bonne NX and Vincent C (2016) Big Data: Coming soon to ENT. *European Annals of Otorhinolaryngology, Head and Neck Diseases* 133(3): 157.
- Robertson H and Travaglia J (2019) Big Data Sociology: Preparing for the Brave New World. In: *The Sociological Review*. Available at: [www.thesociologicalreview.com/big-data-sociology-preparing-for-the-brave-new-world/](http://www.thesociologicalreview.com/big-data-sociology-preparing-for-the-brave-new-world/) (accessed 25 October 2020).
- Rose NS (2007) *Politics of Life Itself: Biomedicine, Power and Subjectivity in the Twenty-first Century*. Princeton: Princeton University Press.
- Rosenstein BS, West CM, Bentzen SM, Alsner J, Nicolaj Andreassen C, Azria D, Barnett GC, Baumann M, Burnet N, Chang-Claude J, Chuang EY, Coles CE, Dekker A, De Ruyck K, De Ruyscher D, Drumea K, Dunning AM, Easton D, Eeles R, Fachal L, Gutiérrez-Enríquez S, Haustermans K, Alberto Henríquez-Hernández L, Imai T, Jones GDD, Kerns SL, Liao Z, Onel K, Ostrer H, Parliament M, Pharoah PDP, Rebbeck TR, Talbot CJ, Thierens Hu, Vega A, Witte JS, Wong P and Zenhausern F (2014) Radiogenomics: Radiobiology Enters the Era of Big Data and Team Science. *International Journal of Radiation Oncology\*Biological\*Physics* 89(4): 709–713.
- Rousseau R (2009) The most influential editorials. *Celebrating scholarly communication studies* 1(1): 47–53.
- Rudinow Sætnan A, Schneider I and Green N (2018) The politics of Big Data: principles, policies, practices. In: Rudinow Sætnan

- A, Schneider I, and Green N (eds) *The politics of Big Data: Big Data, Big Brother?* Abingdon: Routledge, pp. 1-18.
- Rüppel J, Voigt TH (2019) The Death of the Clinic? Emerging Biotechnologies and the Reconfiguration of Mental Health. *Science, Technology, & Human Values* 44(4): 567-580.
- Salerno J, Knoppers BM, Lee LM, Hlaing, WM and Goodman KW (2017) Ethics, Big Data and computing in epidemiology and public health. *Annals of Epidemiology* 27(5): 297-301.
- Sample I (2018) 'It's going to create a revolution': how AI is transforming the NHS. *The Guardian*, 4 July, 18.
- Seaver N (2017) Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society* 4(2): 1-2.
- Shah NH and Tenenbaum JD (2012) The coming age of data-driven medicine: translational bioinformatics' next frontier. *Journal of the American Medical Informatics Association* 19(1): e2-e4.
- Sharon T (2016) The Googlization of health research: from disruptive innovation to disruptive ethics. *Personalized Medicine* 13(6): 563-574.
- Sharon T (2018) When digital health meets digital capitalism, how many common goods are at stake? *Big Data & Society* 5(2): 1-12.
- Sharon T (2020) Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. *Ethics and Information Technology*: 1-13.
- Shelton T (2017) The urban geographical imagination in the age of Big Data. *Big Data & Society* 4(1): 1-14.
- Simon J (2015) Distributed Epistemic Responsibility in a Hyperconnected Era. In: Floridi L (eds) *The Online Manifesto*. Berlin: Springer, pp. 145-159.
- Smits M (2002) *Monsterbezwering*. Amsterdam: Boom Uitgevers Amsterdam.
- Smits M (2006) Taming monsters: The cultural domestication of new technology. *Technology in Society* 28(4): 489-504.
- Somsashekhar SP, Sepúlveda MJ, Norden AD, Rauthan A, Arun K, Patil P, Ethadka RY and Kumar RC (2017) Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. *Journal of Clinical Oncology* 35(15): 8527-8527.
- Spielberg S (2002) *Minority Report*. 20<sup>th</sup> Century Fox, DreamWorks Pictures, Amblin Entertainment, Blue Tulip Productions.
- Star SL and Griesemer JR (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.
- Steeves V (2020) A dialogic analysis of Hello Barbie's conversations with children. *Big Data & Society* 7(1): 1-21.
- Stevens M, Wehrens R and de Bont A (2018) Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society* 5(2): 1-21.
- Stevens M, Wehrens R and de Bont A (2020) Epistemic virtues and data-driven dreams: on similarities and differences between data-driven and medical practices. *Social Science & Medicine* 258(113116): 1-8.
- Stevens M, Wehrens R and de Bont A (Forthcoming) Data-driven dreams: personal sensemaking of data promises for new healthcare information.
- Strauss A and Corbin JM (1990) *Basics of Qualitative Research: Grounded Theory Procedure and Techniques*. London: SAGE.
- Suchman LA (2007) *Human-machine reconfigurations: plans and situated actions*. Cambridge: Cambridge University Press.
- Suchman LA, Follis K and Weber J (2017) Tracking and Targeting: Sociotechnologies of (In) security. *Science, Technology, & Human Values* 42(6): 983-1002.

- Swierstra T (2018) De ethicus als ontregelaar. *Podium voor Bio-ethiek* 25(2): 9-14.
- Tamux, M, Franchois KE and Thomas EJ (2011) What's past is prologue: organizational learning from a serious patient injury. *Safety science* 49(1): 75-82.
- Tanninen M (2020) Contested technology: Social scientific perspectives of behavior-based insurance. *Big Data & Society* 7(2): 1-14.
- Tavory I and Timmermans S (2014) *Abductive analysis: Theorizing qualitative research*. Chicago: University of Chicago Press.
- The Institute for Ethical AI & Machine Learning (2020) The Responsible Machine Learning Principles. Available at: [www.ethical.institute/principles.html](http://www.ethical.institute/principles.html) (accessed 25 October 2020).
- Thrift N (2014) The 'sentient city' and what it may portend. *Big Data & Society* 1(1): 1-21.
- Timmermans S and Berg M (2003) *The Gold Standard: The Challenge of Evidence-based Medicine and Standardization in Health Care*. Philadelphia: Temple University Press.
- Timmermans S and Tavory I (2012) Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory* 30(3): 167-186.
- Törnberg P and Thörnberg A (2018) The limits of computation: A philosophical critique on contemporary Big Data research. *Big Data & Society* 5(2): 1-12.
- Tronto JC (2013) *Caring Democracy: Markets, equality, and justice*. New York: New York University Press.
- Tubaro P, Casilli AA, Coville M (2020) The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society* 7(1): 1-12.
- Tupasela A, Snell K and Tarkkala H (2020) The Nordic data imaginary. *Big Data & Society* 7(1): 1-13.
- Tutton R (2012) Personalizing medicine: futures present and past. *Social Science & Medicine* 75(10): 1721-1728.
- Van Baalen S (2019) *Knowing in Medical Practice: Expertise, Imaging Technologies and Interdisciplinarity*. PhD Thesis, Universiteit Twente, NL.
- Van De Ville D and Lee SW (2012) Brain decoding: Opportunities and challenges for pattern recognition. *Pattern Recognition* 45(6): 2033-2034.
- Van Hulst M and Yanow D (2016) From policy "frames" to "framing" theorizing a more dynamic, political approach. *The American Review of Public Administration* 46(1): 92-112.
- Van Lente H (1993) *Promising Technology: The Dynamics of Expectations in Technological Developments*. Phd Thesis, Universiteit Twente, NL.
- Van Lente H (2012) Navigating foresight in a sea of expectations: lessons from the sociology of expectations. *Technology Analysis & Strategic Management* 24(8): 769-782.
- Vaughan D (1999) The Role of the Organizations in the Production of Techno-Scientific Knowledge. *Social Studies of Science* 29(6): 913-943.
- Venturini T, Jacomy M, Meunier A and Latour B (2017) An unexpected journey: A few lessons from sciences Po médialab's experience. *Big Data & Society* 4(2): 1-11.
- Verbeek PPCC (2014) *Op de vleugels van Icarus: Hoe techniek en moraal met elkaar meebewegen*. Rotterdam: Lemniscaat.
- Verbeek PPCC (2014) Some Misunderstandings About the Moral Significance of Technology. In: Kroes P and Verbeek PPCC (eds) *The Moral Status of Technical Artefacts*. Dordrecht: Springer Science & Business Media, pp. 75-88.
- Verbeek PPCC (2015) Beyond Interaction: A Short Introduction to Mediation Theory. *Interactions* 22(3): 26-31.
- Vidal F and Ortega F (2017) *Being Brains: Making the Cerebral Subject*. New York: Fordham University Press.
- Von Gunten CF, Teno JM and Morrison R (2016) Big Data and End-of-Life Care: Promise

- and Peril. *Journal of Palliative Medicine* 19(12): 1240–1240.
- Wadmann S and Hoeyer K (2018) Dangers of the digital fit: Rethinking seamlessness and social sustainability in data-intensive healthcare. *Big Data & Society* 5(1): 1–13.
- Wainwright SP, Williams C, Michael M, Farsides B and Cribb A (2006) Ethical boundary-work in the embryonic stem cell laboratory. *Sociology of Health & Illness* 28(6): 732–748.
- Wehrens R (2015) The Potential of the Imitation Game Method in Exploring Healthcare Professionals' Understanding of the Lived Experiences and Practical Challenges of Chronically Ill Patients. *Health Care Analysis* 23(3): 253–271.
- Wehrens R, Sihag V, Sülz S, van Elten H, van Raaij E, de Bont A and Weggelaar AM (2020) Understanding the Uptake of Big Data in Health Care: Protocol for a Multinational Mixed-Methods Study. *JMIR Research Protocols* 9(10): 1–9.
- Wehrens R, Stevens M and de Bont A (2019) De ethische framing van Big Data in de zorg. *Podium voor Bio-ethiek* 26(4): 6–9.
- Wehrens R, Stevens M, Kostenzer J, Weggelaar A and de Bont A (Forthcoming) "We can't keep hiding behind the ethics": The role of ethical framing in the promissory discourses of data-driven healthcare.
- Weinstein JN (2016) An 'Industrial Revolution' in Health Care: The Data Tell Us the Time Has Come. *SPINE* 41(1): 1–2.
- Whitman M (2020) "We called that a behavior": The making of institutional data. *Big Data & Society* 7(1): 1–13.
- Wilmott C (2016) Small moments in Spatial Big Data: Calculability, authority and interoperability in everyday mobile mapping. *Big Data & Society* 3(2): 1–16.
- Winter JS and Davidson E (2019) Big Data governance of personal health information and challenges to contextual integrity. *The Information Society* 35(1): 36–51.
- Wouters P and Beaulieu A (2006) Imagining e-science beyond computation In: Hine C (eds) *New infrastructures for knowledge production: Understanding e-science*. Pennsylvania: IGI Global, pp. 48–70.
- Wyatt S (2016) A Computational Turn in the Humanities? A Perspective from Science and Technology Studies. *Journal of Siberian Federal University* 2(9): 517–524.
- Wyatt S, Scharnhorst A, Beaulieu A and Wouters P (2013) Introduction to virtual knowledge. In: Wouters P, Beaulieu A, Scharnhorst A and Wyatt S (eds) *Virtual Knowledge: Experimenting in the Humanities and the Social Sciences*. Cambridge: The MIT Press, pp. 1–23.
- Youtie J, Porter AL and Huang Y (2016) Early social science research about Big Data. *Science and Public Policy* 44(1): 65–74.
- Ziewitz M (2016) Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values* 41(1): 3–16.
- Ziewitz M (2017) A not quite random walk: Experimenting with the ethnomethods of the algorithm. *Big Data & Society* 4(2): 1–13.
- Zinken J, Hellsten I and Nerlich B (2008) Discourse metaphors. In: Frank R, Dirven R, Ziemke T, et al. (eds) *Body, Language, and Mind. Vol. 2: Sociological Situatedness*. Berlin: Mouton, pp. 363–386.
- Zook M, Barocas S, Crawford K, Keller E, Gangadharan SP, Goodman A, Hollander R, Koenig BA, Metcalf J, Narayanan A, Nelson A and Pasquale F (2017) Ten simple rules for responsible Big Data research. *PLoS Computational Biology* 13(3): 1–10.
- Zuiderent-Jerak T (2007) Preventing Implementation: Exploring Interventions with Standardization in Healthcare. *Science as Culture* 16(3): 311–329.

- Zurlinden J (2016) Health Professions Education: Make Big Data Meaningful. *The Journal of Continuing Education in Nursing* 47(5): 199–200.
- Zwitter A (2014) Big Data ethics. *Big Data & Society* 1(2): 1-6.



# Summary





Big Data, machine learning and other data-driven technologies are arriving in healthcare accompanied by hopes and fears. Many of these hopes and fears are based on assumptions about the relation between technology and epistemology. Some people dream that data-driven technology will fundamentally change knowledge practices in healthcare. They hope that data can be used to make better predictions, smarter decisions and to develop personalized interventions. Other people, however, fear that data-driven technologies will lead to incomplete, inscrutable, and misguided decision-making with possibly detrimental effects.

The aim of this dissertation is to critically investigate how data-driven technologies reconfigure responsible knowledge practices in healthcare. Over the years, particular epistemic practices have been developed in healthcare in which norms and values guide what is seen as responsible evidence and information. Data-driven technologies promote the disruption of current epistemic cultures in the healthcare sector. This means that healthcare sector actors must adjust their epistemic practices to these new technologies and determine which data, information and knowledge is responsible to use in healthcare.

This dissertation takes an emic perspective in terms of what is deemed a responsible knowledge practice. It views epistemology as something enacted in practice rather than an external category or a set of abstract criteria. In healthcare, people do not explicitly discuss “epistemology”, but they value particular information, decide what evidence to use and negotiate on methodologies. All these practices give insight into the epistemic cultures of healthcare fields and the norms and values that play a role. The main research question is: *How do various actors reconfigure responsible knowledge practices in data-driven healthcare?*

Like a “scavenger”, I looked around and picked up diverse traces to gain an understanding of how responsible knowledge practices were reshaped, based on data-driven technologies. I needed to move flexibly from site to site, using several methods to gather and analyze the research data. All research data were collected between December 2016 and July 2019 and 164 interviews in total were combined with more than 250 hours of observations and extensive document analyses.

The chapters of this dissertation assemble accounts of diverse epistemic practices in healthcare to which data-driven technologies are introduced, either by envisioning or by day-to-day construction. I trace the envisioning of data-driven healthcare back to writing in scientific journals in the healthcare domain (Chapter 2) and the accounts of diverse experts involved in Big Data initiatives stimulated by the European Union (Chapters 3 and 4). I track down the construction of data-driven technologies to a pioneering hospital-based initia-

tive in the Netherlands (Chapter 5) and a machine learning training program for medical professionals (Chapter 6).

Based on a literature review, **Chapter 2** gives insight into how Big Data is **perceived** in the healthcare sector. We systematically searched the scientific literature, selecting 206 editorials that reflect on developments in the healthcare sector. We describe five discourses constructed on the basis of the editorials. All discourses frame data-driven technologies in particular ways and envision particular consequences. Three discourses (the modernist, instrumentalist and pragmatist) disseminate a compelling rhetoric that presumes that Big Data are benign and lead to valid knowledge. The two other discourses (the scientist and critical-interpretative) question the objectivity and effectivity claims of Big Data but are in the minority. The chapter suggests conducting further research into the influence of such Big Data discourses and the construction of actual Big Data initiatives in healthcare.

**Chapter 3** focuses on the influence of ethical **framing**. It is based on 137 interviews with various experts involved in Big Data initiatives in eight European countries as well as on document analyses. We describe how experts distribute roles and responsibilities regarding ethics in the way they envision data-driven healthcare. The chapter identifies three forms of ethical framing: ethics as a balancing act, a technical fix and a collective thought process. These frames portray the achievement of ethical data-driven healthcare by balancing ethical values, by technical or instrumental problem-solving, or by joint deliberation. This chapter shows that promises and expectations steer toward particular responsible knowledge practices by framing ethics in specific ways.

**Chapter 4** provides deeper insights into why positive discourses seem to prevail. The chapter is based on the same 137 interviews and document analyses mentioned in Chapter 3 but focuses on the epistemic-affective sensemaking process of the experts. We use the metaphor of dreams to describe how experts **dream** that data-driven technologies can help overcome general, scattered, slow and uncontrollable information in healthcare. The chapter also highlights how experts make trade-offs in the way they envision data-driven healthcare, depending on their motivations, values and considerations.

**Chapter 5** is based on six months of ethnographic fieldwork in a Dutch hospital-based data-driven initiative in psychiatry (September 2017 – February 2018). We observe how psychiatrists **negotiate** epistemic differences with data scientists to construct prediction models based on machine learning techniques. We describe how differences were traded locally and temporarily, stimulated by shared epistemic virtues (e.g., systematic approach), boundary objects and socialization processes. Trading became difficult when epistemic virtues were too diverse, differences were enlarged by storytelling and parties did not have

the time or capacity to learn about the other. The chapter ends with a call for bringing epistemic differences into the open as this enables a grounded discussion of the added value of data-driven technology and its role in healthcare.

**Chapter 6** is based on ethnographic fieldwork observing 14 Dutch mental healthcare professionals studying the basics of machine learning on a four-month course while pursuing a machine learning project in their own organizations. The chapter draws upon feminist literature on care to study how both the technology and the medical professionals **care** for responsible knowledge production in their own ways and how epistemic responsibility is shared in a network. We argue that this care perspective reveals patterns in the kind of responsibility taken and points to opportunities to avoid neglect and irresponsibility.

**Chapter 7** presents the conclusions of this research and **reflects** on the responsible knowledge practices that arise in response to data-driven technologies. I conclude that contrary to modernist visions, data-driven healthcare will not lead to big changes in knowledge practices in the short term. It is a slow process in which collaboration between actors is crucial.

The chapters show that data-driven technologies are by no means introduced to an empty playing field but arrive in settings in which all sorts of epistemic cultures, norms and values continue to play a role. Data-driven technologies, data scientists and others are added to the complex knowledge-producing networks, where discussions and negotiations are taking place about epistemic differences and how responsible data-driven healthcare should be shaped.

Three points emerge from the research that are currently crucial for organizing responsible knowledge practices with the new technologies. First, data-driven initiatives should be human-centered. This means that human supervision and control remain essential conditions. Second, the epistemic practices of the healthcare field are leading. This means that healthcare professionals are reluctant to change their epistemic practices too abruptly and hold onto many elements of their epistemic cultures. Third, data-driven technologies should be used in certain healthcare settings. These settings could be (1) deemed safe and manageable as they have clear boundaries, or (2) are currently known to be highly uncertain, meaning that there is little information or evidence available to base decisions on.

This dissertation contributes in three important ways to current scholarly work. First, it approaches ethics as an empirical phenomenon. This perspective shows more precisely what norms and values are at stake in situ. Second, it centralizes epistemology-in-practice.

This offers a nuanced view that picks up on small changes in epistemic practices and differences in epistemic cultures. Lastly, it puts forward six concepts that can be used to study responsible knowledge practices in healthcare.

This dissertation has implications for healthcare practice, with its ongoing negotiations on responsible knowledge practices and data-driven technologies. First, healthcare actors tend to overestimate the pace of changes in healthcare. There is time for reflection on epistemic practices. Second, it is important to make data-driven initiatives resilient by focusing on how responsible knowledge practices are created and by learning from examples. Third, continuity is important as data-driven technologies need constant adaptation to the healthcare practices in which they are used and vice versa. Fourth, appropriate rules are necessary to ensure that negotiations about epistemic practices can proceed and data-driven initiatives are safeguarded from too large power asymmetries.

Social scientists and philosophers have an important role in pro-actively and continuously following the construction of data-driven healthcare. Much uncertainty surrounds data-driven technologies. I stress the importance of doing empirical research and continuously following technological developments and the people, data, methods, cultures, norms and values in healthcare practices. These insights can stimulate discussions in the field, bring in new perspectives and help develop responsible knowledge practices with data-driven technologies.





# Samenvatting





Big data, data science, machine learning en andere data-gedreven technologieën worden in de gezondheidszorg ontvangen met hoop en angst. In veel van de reacties worden aannames gemaakt over de relatie tussen de technologieën en epistemologie. Sommige mensen dromen dat data-gedreven technologieën de kennispraktijken in de gezondheidszorg fundamenteel zullen veranderen. Ze hopen dat data gebruikt kunnen worden om betere voorspellingen te doen, slimmere beslissingen te nemen en gepersonaliseerde interventies te ontwikkelen. Andere mensen vrezen echter dat data-gedreven technologieën zullen leiden tot onvolledige, ondoorgrondelijke en misleidende besluitvorming met mogelijk schadelijke gevolgen.

Het doel van dit proefschrift is om kritisch te onderzoeken hoe data-gedreven technologieën verantwoorde kennispraktijken in de gezondheidszorg veranderen. In de loop der jaren zijn in de gezondheidszorg specifieke kennispraktijken ontwikkeld, waarin bepaalde normen en waarden richtinggevend zijn voor wat als verantwoord bewijs en informatie wordt gezien. De verwachting is dat data-gedreven technologieën gaan leiden tot veranderingen in de bestaande kennispraktijken. Actoren in de gezondheidszorg moeten zich verhouden tot deze nieuwe technologieën en beslissen hoe ze op een verantwoorde manier de nieuwe data, methodologieën en informatie kunnen gebruiken.

Dit proefschrift gebruikt een emic perspectief om te kijken naar verantwoorde kennispraktijken. Dat betekent dat epistemologie niet als een externe categorie of een reeks abstracte criteria wordt gezien, maar als iets dat in de praktijk wordt gedaan. In de gezondheidszorg spreken mensen niet expliciet over “epistemologie”, maar toch waarderen ze bepaalde informatie, maken ze beslissingen over welk bewijs ze gebruiken en onderhandelen ze over methodologieën. Al deze acties geven inzicht in de kennispraktijken in de gezondheidszorg en welke normen en waarden daarbij een rol spelen. De hoofdvraag van dit onderzoek is: *hoe veranderen diverse actoren verantwoorde kennispraktijken in data-gedreven gezondheidszorg?*

Om inzicht te krijgen in de veranderende kennispraktijken ben ik als een “aaseter” onderzoek gaan doen. Ik keek rond, zocht verschillende sporen om een idee te krijgen hoe verantwoorde kennispraktijken veranderen door data-gedreven technologieën. Ik heb me flexibel door verschillende omgevingen bewogen en heb diverse methoden gebruikt om mijn onderzoeksgegevens te verzamelen en te analyseren. Alle onderzoeksgegevens zijn verzameld tussen december 2016 en juli 2019. In het totaal heb ik 164 interviews gecombineerd met meer dan 250 uur aan observaties en uitgebreide documenten analyses.

In de hoofdstukken van dit proefschrift rapporteer ik over diverse praktijken in de gezondheidszorg waarin data-gedreven technologieën worden geïntroduceerd. Dit kan

zijn door middel van verwachtingen of in daadwerkelijke initiatieven. Ik onderzoek de verwachtingen van data-gedreven technologieën in wetenschappelijke tijdschriften uit de gezondheidszorg (Hoofdstuk 2) en in de verhalen van experts die betrokken zijn bij Big Data initiatieven die gestimuleerd zijn door de Europese Unie (Hoofdstukken 3 en 4). Daarnaast onderzoek ik de constructie van data-gedreven gezondheidszorg in een innovatief ziekenhuis initiatief in Nederland (Hoofdstuk 5) en een machine learning opleiding voor medische professionals (Hoofdstuk 6).

**Hoofdstuk 2** is gebaseerd op een literatuurstudie en geeft inzicht in hoe Big Data wordt ontvangen in de zorg. We hebben systematisch de wetenschappelijke literatuur doorzocht en 206 editorials geselecteerd die reflecteren op de ontwikkelingen in de zorgsector. We beschrijven 5 discoursen die we hebben geconstrueerd op basis van de editorials. Elk discours beschouwt data-gedreven technologieën op een bepaalde manier en voorziet specifieke gevolgen. Drie discoursen (het modernistisch, instrumentele en pragmatische) verspreiden een heel overtuigende retoriek die veronderstelt dat Big Data goed zijn en tot geldige kennis leiden. Twee andere discoursen (het wetenschappelijke en kritisch-interpretatieve) stellen de objectiviteit en effectiviteit van Big Data ter discussie, maar zijn in de minderheid. Het hoofdstuk raadt verder onderzoek naar de invloed van Big Data discoursen en de daadwerkelijk Big Data initiatieven in de zorg aan.

**Hoofdstuk 3** richt zich op de invloed van ethische framing. Het hoofdstuk is gebaseerd op 137 interviews met diverse experts die betrokken zijn bij Big Data-initiatieven in acht Europese landen in combinatie met documentanalyses. We beschrijven hoe experts rollen en verantwoordelijkheden met betrekking tot ethiek verdelen in de manier waarop zij zich data-gedreven gezondheidszorg voorstellen. Het hoofdstuk identificeert drie vormen van ethische framing: ethiek als evenwichtsoefening, een technische fix en een collectief denkproces. Volgens deze frames is ethische data-gestuurde gezondheidszorg te bereiken door respectievelijk bepaalde ethische waarden in evenwicht te brengen, door technische of instrumentele probleemoplossing of door gezamenlijk overleg. Dit hoofdstuk laat zien dat beloften en verwachtingen aansturen op bepaalde verantwoorde kennispraktijken door ethiek op een specifieke manier te kaderen.

**Hoofdstuk 4** geeft meer inzicht in waarom positieve discoursen zo dominant lijken te zijn. Het hoofdstuk is gebaseerd op dezelfde 137 interviews en documentanalyses als hoofdstuk 3, maar richt zich op de betekenisgeving van de experts. We gebruiken de metafoor van dromen om te beschrijven hoe experts dromen dat data-gedreven technologieën kunnen helpen algemene, verspreide, langzame en oncontroleerbare informatie in de gezondheidszorg te verbeteren. Het hoofdstuk laat ook zien hoe experts bepaalde

afwegingen maken in de manier waarop ze data-gedreven gezondheidszorg voorstellen, afhankelijk van persoonlijke motivaties, waarden en overwegingen.

**Hoofdstuk 5** is gebaseerd op zes maanden etnografisch onderzoek in een Nederlands data-gedreven initiatief in de psychiatrie (september 2017 - februari 2018). We beschrijven hoe psychiaters onderhandelen over epistemische verschillen met data scientists om voorspellingsmodellen op basis van machine learning-technieken te ontwikkelen. De verschillen werden lokaal en tijdelijk uit onderhandeld en gestimuleerd door gedeelde epistemische deugden (zoals een systematische benadering), “boundary” objecten en socialisatieprocessen. De onderhandelingen werden moeilijk als deugden te divers waren, verschillen vergroot werden door verhalen die gingen rondzingen of als partijen niet de tijd of capaciteit hadden om de ander goed te leren kennen. Het hoofdstuk sluit af met een oproep om epistemische verschillen beter te onderzoeken, aangezien dit een genuanceerde discussie mogelijk maakt over de meerwaarde van data-gedreven technieken en hun rol in de gezondheidszorg.

**Hoofdstuk 6** is gebaseerd op etnografisch veldwerk waarbij 14 Nederlandse professionals in de geestelijke gezondheidszorg worden geobserveerd die de basisprincipes van machine learning leren tijdens een cursus van vier maanden. Als onderdeel van de cursus moesten de professionals ook een machine learning-project opzetten in hun eigen organisatie. Het hoofdstuk maakt gebruik van feministische literatuur over “care” om te bestuderen hoe zowel de technologie als de medische professionals op hun eigen manier zorgen voor verantwoorde kennisproductie en hoe epistemische verantwoordelijkheid wordt gedeeld in een netwerk. Wij stellen dat het zorg perspectief helpt om de verschillende manieren waarop verantwoordelijkheid genomen wordt te bestuderen en kan helpen om verwaarlozing en onverantwoordelijkheid te voorkomen.

**Hoofdstuk 7** presenteert de conclusies van dit onderzoek en reflecteert op de verantwoorde kennispraktijken die ontstaan in reactie op data-gedreven technologieën. Ik concludeer dat in tegenstelling tot modernistische visies, data-gedreven zorg niet op korte termijn tot grootschalige verandering van kennispraktijken gaat leiden. Het is een langzaam proces waarbij samenwerking tussen actoren cruciaal is.

De hoofdstukken laten zien dat data-gedreven technologieën geenszins in een leeg speelveld worden geïntroduceerd, maar terechtkomen in een omgeving waarin allerlei epistemische culturen, normen en waarden een rol blijven spelen. Data-gedreven technologieën, data scientists en anderen worden toegevoegd aan de complexe netwerken waarin kennis wordt geproduceerd. In dergelijke netwerken vinden discussies en onderhandelingen plaats over epistemische verschillen en hoe verantwoorde data-gedreven zorg vorm moet krijgen.

Uit het onderzoek komen drie punten naar voren die op dit moment cruciaal zijn voor het organiseren van verantwoorde kennispraktijken met de nieuwe technieken. Ten eerste is een mensgerichte benadering belangrijk. Dit betekent dat er data-gedreven initiatieven gekozen worden voor menselijk toezicht en controle. Ten tweede blijken de epistemische culturen van de gezondheidszorg leidend. Dit betekent dat zorgprofessionals terughoudend zijn om hun kennispraktijken abrupt te veranderen en blijven zij vasthouden aan veel elementen uit hun eigen cultuur. Ten derde mogen data-gedreven technologieën alleen gebruikt worden voor bepaalde vraagstukken. Dit kunnen vraagstukken zijn die (1) als veilig worden beschouwd omdat ze duidelijk afgrensbaar en beheersbaar zijn of (2) momenteel bekend staan als zeer onzeker, wat betekent dat er weinig informatie is of bewijs om beslissingen op te baseren.

Het gekozen perspectief van dit proefschrift draagt op drie manieren bij aan het huidige wetenschappelijke onderzoek. Ten eerste benadert het proefschrift ethiek als een empirisch fenomeen. Dit perspectief laat nauwkeuriger zien welke normen en waarden er in de praktijk op het spel staan. Ten tweede centraliseert het epistemologie-in-de-praktijk. Dit biedt een genuanceerd beeld dat rekening houdt met kleine veranderingen in kennispraktijken en verschillen in epistemische culturen. Ten slotte voegt het toe door zes concepten te presenteren die kunnen worden gebruikt om verantwoorde kennispraktijken in de gezondheidszorg te bestuderen.

Dit proefschrift heeft implicaties voor de zorgpraktijk, waarin de onderhandelingen over verantwoorde kennispraktijken en data-gedreven technologieën gaande zijn. Ten eerste hebben zorgactoren de neiging om het tempo van veranderingen in de gezondheidszorg te overschatten. Dit proefschrift laat zien dat er tijd is om stil te staan bij de epistemische veranderingen. Ten tweede is het belangrijk om data-gedreven initiatieven veerkrachtiger te maken door te focussen op de manieren waarop verantwoorde kennispraktijken worden gecreëerd en door te leren van goede voorbeelden. Ten derde is continuïteit belangrijk in data-gedreven initiatieven, aangezien data-gedreven technologieën voortdurend moeten worden aangepast aan de zorgpraktijken waarin ze worden gebruikt en vice versa. Ten vierde zijn er passende regels nodig om ervoor te zorgen dat onderhandelingen over kennispraktijken door kunnen gaan en data-gedreven initiatieven worden beschermd tegen te grote machtsasymmetrieën.

Sociale wetenschappers en filosofen spelen een belangrijke rol bij het proactief en continu volgen van data-gedreven initiatieven in de zorg. Er is op dit moment veel onzekerheid over hoe data-gedreven technologieën gebruikt moeten worden. Ik benadruk het belang van empirisch onderzoek, het continu volgen van technologische ontwikkelingen en de mensen, data, methoden, culturen, normen en waarden in de zorgpraktijk. Deze inzichten

kunnen discussies in het veld stimuleren, nieuwe perspectieven inbrengen en helpen bij het ontwikkelen van verantwoorde kennispraktijken met data-gedreven technologieën.



# Curriculum Vitae





## PhD Portfolio

Name: Marthe Stevens

Department: Erasmus School of Health Policy & Management

PhD period: 2016 – 2020

Promotor: Prof. dr. Antoinette de Bont

Co-promotor: dr. Rik Wehrens

### Courses

EGSH	Project management for PhDs	2016
EGSH	Doing the literature review	2017
WTMC workshop	Art	2017
EGSH	Ethnography and its varieties	2017
Erasmus MC	Diagnostic data for dummies	2017
WTMC summer school	Digital ethnography	2017
EIT Health summer school	Design Thinking	2017
WTMC workshop	Responsible Research and Innovation	2017
Nobbe Mieras	Speak up dear	2018
WTMC workshop	Doing comparison	2018
Maastricht University	Skillshare writing workshop	2018
WTMC summer school	Infrastructures	2018
EGSH	Philosophy of the social sciences and humanities	2018
Risbo	Basic didactics	2019
Risbo	Group dynamics	2019
WTMC workshop	Post-Colonial	2019
Rathenau Institute	Providing evidence for policy making	2020
EGSH	Self-presentation: focus, structure and interaction	2020
Risbo	University Teaching Qualification (UTQ)	2020

### Peer-reviewed publications

Stevens M, Wehrens R and de Bont A (2018) Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society* 5(2): 1-21.

Stevens M, Wehrens R and de Bont A (2020) Epistemic virtues and data-driven dreams: on similarities and differences between data-driven and medical practices. *Social Science & Medicine* 258, 1-8.

Wehrens R, Stevens M and de Bont A (2019) De ethische framing van Big Data in de zorg. *Podium voor Bio-ethiek* 26(4): 6-9.

Stevens M and Beaulieu A. Relating to epistemic responsibility: the machine learning entanglements of ten mental healthcare organizations in the Netherlands. *Under review at Science Technology, & Human Values*.

Stevens M, Wehrens R and de Bont A. Data-driven dreams: personal sensemaking of data promises for new healthcare information. *Under review at Big Data & Society*.

Wehrens R, Stevens M, Kostenzer J, Weggelaar-Jansen AM and de Bont A. - "We can't keep hiding behind the ethics": The role of ethical framing in the promissory future of data-driven healthcare technologies. *Under review at Science Technology, & Human Values*.

Felder M, Kleinhout-Vliek T, Stevens M and de Bont A. Design-thinking as a ritualized change strategy in the governance of care: an ethnographic case study. *Submitted to Policy Sciences*.

Wehrens, R, Stevens M, Kostenzer J, Weggelaar-Janssen A and de Bont A. Data-driven healthcare and the quest for legitimacy: an international comparison of regulatory approaches. *To be submitted*.

## Other publications

Stevens M (2019) *Evaluatierapport: Opleiding Machine Learning voor de Klinische Praktijk*. Erasmus University Rotterdam.

Wehrens R, Stevens M, Kostenzer J, Weggelaar-Janssen A and de Bont A (2019) *Infographics on regulations for big data technology in the healthcare sector in European Countries*.

## Presentations

- |            |  |
|------------|--|
| 30/08/2017 | <i>Big data: beyond the talk?</i><br>Annual meeting of the Society for Social Studies of Science (4S), Boston  |
| 02/05/2018 | <i>Dating with monsters: the practices of Big Data in healthcare</i><br>WTMC workshop, Ravenstein  |
| 30/11/2018 | Discussant of the book "Platform Society" (invited talk)<br>Annual Meeting WTMC, De Balie, Amsterdam   |
| 12/12/2018 | <i>Dating with monsters: Big Data in a psychiatric hospital department</i><br>IFIP 8.2. Working conference: Living with monsters? San Francisco State University |

- 21/03/2019 *Big Data in a psychiatric hospital department: trading epistemic virtues*  
Big Data and the Power of the Narrative Seminar, IT University of Copenhagen
- 08/04/2019 *Big Data en de psychiatrie*  
UMC Utrecht
- 11/06/2019 *Bevindingen: Opleiding Machine Learning voor Medisch Professionals*  
Trimbos Institute, Utrecht
- 12/03/2020 *Epistemic virtues and data-driven dreams*  
MultiJuSe event, Erasmus University Rotterdam
- 07/12/2020 *Reflections on data-driven mental healthcare* (invited talk)  
SIG Clinical Data – Circulatory Health Meeting, Utrecht University

## Teaching activities

### *Bachelor Health Sciences, Erasmus University Rotterdam*

- 2016 - 2018 Workgroups "Policy and management studies"
- 2017 Workgroups "Qualitative research methods"
- 2018 - 2019 Thesis supervision

### *Master Health Care Management, Erasmus University Rotterdam*

- 2016 - 2020 Workgroups "Health Care Governance"
- 2018 - 2020 Co-developer, co-coordinator, lecturer and workgroups in "The potential and challenges of Big Data"
- 2019 - 2020 Thesis supervision

### *Guest lectures*

- 2019 - 2020 Guest lecturer, Managing Science and Technology, VU Amsterdam
- 2020 Guest lecturer, Governing Healthy Cities, Erasmus University College

## Research projects

09/2017 – 02/2018

Ethnographic research during a design thinking workshop for the Dutch healthcare inspectorate, a health insurer and healthcare provider, interviews with participants, analysis of the results.

09/2017 – 02/2018

Ethnographic research during the development of prediction models for psychiatric diseases and interviews with stakeholders at a large Dutch Academic Hospital, analysis of the results.

05/2018 – 12/2018

Interviews with stakeholders in Spain and Germany for a cross-country comparison of big data governance, analysis of results and development of infographics. All for the Horizon 2020 project “Big Medilytics”.

03/2019 – 07/2019

Ethnographic research during a machine learning training for healthcare professionals and interviews at the Trimbos Institute (the Netherlands), analysis of the results and writing evaluation report for ZonMw.

### **Awards and nominations**

- |      |  |
|------|--|
| 2019 | Nominated Big Data Value Success Story Award<br>BDV PPP Summit, Riga   |
| 2021 | Best Article 2020, Graduate School Award for PhD Excellence<br>Graduate School of Social Science and the Humanities, Rotterdam |

### **Additional activities**

- |             |   |
|-------------|---|
| 2017 - 2019 | Board member of PhD Council young-ESPHM                             |
| 2017 - 2019 | Organizer of the annual PhD survey about well-being and supervision |
| 2017 - 2018 | Supervision junior researcher (Mark Zijdemans)                      |

## About the author

Marthe Stevens (1990) studied health sciences and philosophy at the Erasmus University of Rotterdam. During her studies, she was a (board) member of various university and faculty councils. Marthe did internships at the Rathenau Institute and Utrecht Medical Centre and obtained her language proficiency certificates in London. She also worked as a research assistant helping with an evaluation study for community-dwelling frail older persons and teaching bachelor students quantitative research skills.

After her studies, she started with her PhD research in the field of Science and Technology Studies at the Erasmus School of Health Policy and Management. She worked together with a diverse set of Dutch (mental) healthcare organizations, participated in the EU-funded BigMedilytics project and design thinking projects and collaborated with Anne Beaulieu from the University of Groningen. During her PhD research, she was active in the PhD representation board of the Erasmus School of Health Policy and Management. In addition, she was a tutor and lecturer in various courses and taught to diverse groups of bachelor and master students. She co-founded the master course “the potential and challenges of Big Data” for healthcare management students.

Currently, Marthe is a postdoctoral researcher at the Faculty of Philosophy, Theology and Religious Studies at Radboud University Nijmegen. She works at the Interdisciplinary Hub for Security, Privacy and Data Governance (iHub) on the ERC-funded project “The Digital Good” led by dr. Tamar Sharon. Marthe continues to live in Rotterdam with her husband.

For more information: [www.marthestevens.com](http://www.marthestevens.com)

Big data and machine learning technologies often come with dreams to fundamentally change knowledge practices in healthcare for the better. This book asks what happens when such “data-driven dreams” become drivers for concrete initiatives in healthcare. How are knowledge practices reshaped through the dreams and practices of big data and machine technologies? Marthe Stevens uses a combination of qualitative research methods to shed light on this question by studying how various actors and organizations envision and construct data-driven healthcare. This book shows that current data-driven initiatives are less radically innovative than often assumed. Healthcare organizations and professionals need big dreams to take incremental steps in improving knowledge practices. Here lies the importance of the connection between dreams and practices – they prompt the necessary reflection on the knowledge practices in healthcare. The book will be of interest to anyone seeking to better understand the profound ways in which data-driven technologies affect knowledge practices in health care: scholars, professionals and policymakers alike.

# Dreaming with data

Marthe Stevens