# 1 DATA PROMISES

*October 2017:* "We need to disrupt all value chains!" It is said with conviction, and it comes from a high-ranking civil servant in the Danish healthcare system. The feeling of being given an order makes me straighten up in my chair in the audience. Without explaining what "all" these "value chains" are, nor what this "disruption" will involve in practice, she continues: "We have the most amazing data resources in Denmark, but if we do not act now, we will fall behind!" She is speaking at an annual meeting called the Health Observatory, a gathering of approximately 500 people working with digital health in Denmark. She recounts with agitation and delight what she has learned from a study trip to the United States (and, it seems, particularly from speaking to "a guy doing those algorithms for Amazon and Uber"). It is the third talk during that conference where the speaker mentions a study trip to the United States, and I am feeling increasingly puzzled. I cannot follow how her take-home messages logically connect. I am aware that though the setup looks like an academic conference, talks at this type of meeting adhere to different criteria of success than research presentations, but still: How do "amazing data resources" become an obligation to "disrupt"? I am also beginning to wonder: Who are they visiting on these study trips, and why?

It is not my first time attending these Health Observatory meetings—typically referred to by the participants as "family gatherings" because it is more or less the same group of people coming every year (see also Lindholm and Lerche 2019). I love the welcoming atmosphere, as well as the frank and forthright introduction to what is going on in digital health in Denmark. During lunch later that day, I learn that several of the civil servants who are

attending the meeting have taken courses at Singularity University in Silicon Valley. One man says with profound gravitas as he is leaning over the table toward me: "The world will never look the same to you again!" Apparently, these trips help the prophecies of digital disruption travel from circles in Silicon Valley to Danish hospital wards and administrative corridors. Singularity University also had opened a branch in Denmark just months earlier that year (called SingularityU because "university" is a legally defined name in Denmark that can be used only by those acknowledged by the authorities as such). Study trips to the United States are not the only recurring topic in the talks this year. The very same five PowerPoint slides appear in several talks. These slides were even sent to me by a kind informant a few weeks earlier. They are used to explain the investment of half a billion Danish kroner (DKK) into something called "the health data program." They represent some of the recent investments that were done, so that Denmark does not fall behind.

When I first received these slides, I was thrilled! I thought they contained just what I needed to know about ongoing initiatives aimed at intensified data sourcing in Danish healthcare. The slides mention a 2013 Organisation for Economic Co-operation and Development (OECD) report in which it is pointed out that "the goldmine [of Danish data] is only partly exploited" (OECD 2013: 18). Danish hospitals, the report observes, collect many data that are left unused. The next slide explains that, based on this report, the government asked the consultancy group Deloitte to advise on how to optimize health data use. Deloitte (2014) suggested centralizing data flows to facilitate easier data reuse. The consultancy described the overall purpose as a matter of "supporting a culture in the health services, where the point of departure for medical and health economic actions is *based on data and knowledge about what works*" (6, emphasis added).

In the years that follow, I realize that the Health Data Program was just one investment among many. The slides did not *explain* the ongoing intensification of data sourcing in Denmark. They represented just one particular assemblage of consultancy reports, investments, and perceptions of data needs that all feed into the same overall impetus toward intensified data sourcing. Even the talk about the "need to disrupt all value chains," which first confused me, begins to make sense when seen as part of this wider discourse. The various talks convey the same eschatological sense of being on the verge of something both frightening and attractive. These consultancy reports and talks embody potent *data promises*.

In this chapter, I explore data promises and how they interact with data infrastructures on the ground. Data promises are paradoxical drivers of intensified data sourcing. They are justified by a need for—and promise a future of—decisions based on some sort of "evidence," *but they are not themselves supported by evidence*. First, I reflect on existing studies of promissory politics and how these studies can help us understand data promises as they come across in reports and strategy papers. I then turn to how local infrastructures and political histories interact with global developments by turning to the actors requesting more data for four overlapping purposes: research, clinical use, governance, and economic growth.

## PROMISES: THE WORK OF CONSULTANCY REPORTS AND STRATEGY PAPERS

Promises lie at the heart of contemporary data politics. Sociologist Alan Petersen (2019) suggests that digital health is consistently promoted through promissory discourses where statements about what *will* occur become arguments for not only letting it occur, but for *investing* in making it happen. Callahan (2009) finds in such techno-optimistic promises a motor for escalating healthcare costs. Science and technology studies (STS) scholars have had a long-standing interest in the social lives of technological promises and expectations (Koch 2006; Rapp 2011). They have pointed out that though expectations announce the future, their political power operates in the present (Hedgecoe 2004; Brown, Kraft, and Martin 2006; Brown and Michael 2003; Vezyridis and Timmons 2021). While some dismiss optimistic promises as hype and call for more realistic expectations (Löfgren and Webster 2019), STS has warned against searching for the truth behind the buzz (Vincent 2014). The most urgent task is to explore what promises to do—here and now—in the environments where they circulate. This is also my goal when exploring how promissory policies interact with practice and experience.

As stated in the introduction, data promises generate investments in infrastructure. Investments in infrastructure "signal the desires, hopes, and aspirations of a society, or of its leaders" (Appel, Anand, and Gupta 2018: 19). These signals can easily overrule any documented need for such infrastructures. They depend on narratives and metaphors, and when effective, Annas (2014) argues that they "can make even out of control and extraordinarily expensive quests seem much more reasonable and supportable in theory than they

are likely to be in practice" (226). Promises can be too big to fail, as it were, in the sense that they raise expectations that cannot be cogently evaluated in the same way as everyday requests for financial support for more limited projects. Therefore, big infrastructural investments are sometimes more likely to receive funding than small projects; or, rather, they acquire their funding based on rather fluffy claims and arguments (Davies, Frow, and Leonelli 2013). Morrison argues that because all innovation by definition springs from an infrastructure, it has turned out to be rhetorically effective—though not logically necessary—to claim that investments in infrastructure will lead to innovation (Morrison 2017). Lee (2015) suggests that infrastructural investments are also politically appealing because different actors can align them with very different hopes. Infrastructural investments carry great appeal.

With metaphors for data such as the new oil and gold mines (NordForsk 2014; Palmgren 2017), it should not seem strange that data promises tend to revolve around profit and national growth (Vezyridis and Timmons 2021). Consultancies are often commissioned to quantify a potential profit: they tend to offer a number. Their estimates typically evade academic standards of transparency, but such numbers can make thoughts about a future return on infrastructural investments acquire an aura of factuality (Merry 2016; Espeland and Stevens 2008).

To give an example, I have several times heard policymakers at conferences and during informal meetings reference a particular report from McKinsey when arguing the need to invest in digital data infrastructures in healthcare (McKinsey & Company 2016). The report claims that Sweden would be able to cut its healthcare expenditure by 25 percent through digitization. In its argumentation, it makes a comparison to Denmark—as Sweden's "more digitally advanced" neighbor—although healthcare in Denmark is not 25 percent cheaper. The report has a brief "Methods" section that simply says that the claims are based on a reading of the literature (yet it provides no references), on statistics (but it does not specify which ones), and on 100 interviews (without clarifying with which types of actors or providing any quotes). It is all the more curious that *Danish* policymakers mention the report as proof that Denmark could save 25 percent through digitalization, thereby disregarding the fact that according to the report, Denmark should already have achieved the savings (in a period of digitization when healthcare costs went up). In short, it is a report with numbers that provide rhetorical strength, but with no clear evidence to support the claims it is

used to propagate. Still, most policymakers know how such consultancies work. When I complained to a civil servant about the lack of documentation of the 25 percent (in the course of a meeting where it came up), she laughed and said, "You know, we always say, 'if you need a higher number, then call McKinsey.'"

Data promises encourage investments for fear that the country will fall behind. Speed consistently comes across as a necessary parameter (Jackson 2017). The recurring articulation of "falling behind" that I also heard at the conference with which I opened this chapter has made me wonder: Behind *whom*? In most talks and reports, the nature of the competition remains blurred (Digitaliseringspartnerskabet 2021; Regeringen 2021). Although people go on study trips to the United States and marvel at the growth of its big tech companies, few people consider the US healthcare system worth copying. It is certainly not leading with respect to data integration. The same people who return with a sense of thrill from their US study trips typically remark that American healthcare is scattered and expensive and has left data integration to data brokers. In a 2016 report called "Digitizing Denmark. How Denmark Can Drive and Benefit from an Accelerated Digitized Economy in Europe," commissioned by Google and written by Boston Consulting Group, I finally find some competitors mentioned by name:

> Looking outside of Europe, several Asian countries (Hong Kong, China, Taiwan, Singapore, and South Korea) are highly digitized and/or are undergoing rapid digitization. There is a risk of Denmark being rapidly surpassed by these more digitally proactive economies, leaving the nation in a digital backwater on the global scene, with capital, talent, and growth being focused elsewhere. The value at stake for Denmark . . . translates into a potential 150,000 net full-time equivalent (FTE) positions and more than 200 billion DKK added to the GDP by 2020, an 83% increase in the GDP growth rate. (Alm, Colliander, Gotteberg, et al. 2016a, 4)

"Digital backwater" does not sound very nice, and although the calculation of FTE positions remains opaque, it sounds scary. However, it also sounds like a very generic description of economic threats. The narrative of falling behind Asia has been around since the 1980s (Deming 1986). There are no references to support the claim. Singapore is highly digitized, and yes, South Korea is too, and China is investing in an integrated tracking system, but the rest of Asia is far below the Danish level of digitization. Furthermore, how will competition in this form of digitization between countries affect an 83 percent increase in the gross domestic product (GDP) growth rate?

Numbers again seem to serve as rhetorical devices that give effect to the narrative, rather than as evidence in an academic sense. The report sums up its recommendations as follows:

> Denmark must make faster and broader digitization a national top priority. It is essential for the country in order to secure future GDP growth and create new jobs, as well as to stay competitive in a global and increasingly digital world. (3)

This conclusion lays the groundwork for a nice market for information and communications technology (ICT) companies like the one financing the report. I cannot imagine how digitalization could be given higher priority than it already has, but the report insists that "being in a good position today does not mean Denmark can relax its efforts" (3). Still, the Boston Consulting Group is firm enough in its convictions to have made a whole series of *very similar* reports. One of them carries the title "Digitizing the Netherlands" (and the same subtitle as the one on Denmark). It notes—with a strikingly similar narrative—the following:

> The Netherlands, as one of the leading digital frontrunner nations in Europe, must make further digitization a top priority . . . to secure growth and jobs in a rapidly changing digital world. (Alm et al. 2016: 3)

Yet another report called *Digitizing Europe*—surprise!—recommends similar investments for all of Europe. Much as Ferguson (1994) once noted with respect to policies for development aid, the framing of the problem and solution often remains the same, regardless of context. It might incite us to view the Boston Consulting Group as selling *data promises* with this report; not "data and knowledge about what works" (to use the phrase from Deloitte). Consultancy reports seem to be relatively free to make such authoritative "recommendations" without documentation, at least when they serve already powerful agendas. Pollock and Williams (2010) use the term "promissory organisations" for consultancies that are themselves relieved from keeping their promises. When policy priorities change, they can write a new report.[1]

Data promises are fed not only by consultancies, but also by computer scientists. Domingos (2015), for example, talks about a Master Algorithm that designs itself and cuts itself loose from human interference, an algorithm that

> can derive all knowledge in the world—past, present, and future—from data. Inventing it would be one of the greatest advances in the history of science. . . . The Master Algorithm is our gateway to solving some of the hardest problems we face, from building domestic robots to curing cancer (xviii).

Computer scientists discuss this idea because they already now work with algorithms that develop other algorithms, so that humans do not know how the latter algorithms work (Tegmark 2017). Because digital data analysis is associated with great power, data promises also provoke warnings. Hope and fear go hand in hand (Mulkay 1993). Some computer scientists now fear that such a Master Algorithm will take control of all life (Tegmark 2017). They fear a digital power that will erase the human world as we know it. The sociologist Richard Tutton observes how the old utopian vision of progress more broadly has acquired a dystopian outlook in the inner circles of Silicon Valley, where private investments are now made to facilitate escape from the Earth to settle on Mars following potential disasters (Tutton 2017; Tutton 2020). In interesting ways, the dystopian scare just adds to the gravitas of the digital challenge and, for many of the people I have met, underlines the pivotal importance of being part of the "data-driven" future. If a Master Algorithm is about to take over, it is better to be prepared. Fear can also be a mobilizing promise.

The politics of data promises may fly on the wings of eschatological gospel, but the changes they install strike real people in the present. The implications are often a lot more mundane and low-tech—less sophisticated and less accurate than the consultancy promises suggest. This is why I keep insisting on relating the policy visions to the practices and experiences that patients and health professionals have with contemporary healthcare. For patients, the quest for evidence is much more than a rhetorical game or a business opportunity (Rabeharisoa, Moreira, and Akrich 2014; Rabeharisoa, Callon, Filipe et al. 2014). They are confronted with suffering and despair (Novas 2006; Petersen and Tanner 2015). It is because of them that we need to keep questioning which types of investments data promises stimulate. Patients depend on responsive healthcare systems (Lomborg, Langstrup, and Andersen 2020). Who, then, are the people wanting more data, and what do they wish to do with them? Why do they believe that healthcare needs more data?

## DATA MULTIPLE: MANY ACTORS, MANY PURPOSES

To understand the drivers for intensified data sourcing, we must understand why local decision makers think they need more data. Regardless of the problem, global gospel does not rule the world simply by sounding

exciting. Promises must speak to local concerns and perceptions of problems (Kieser 1997; Røvik 1996). At the same time, nobody just "needs" data. An urge for data is not like thirst or hunger. It is an acquired taste—one nurtured through university training and organizational interactions. So what have people learned to desire?

I have compiled the most common purposes that are said to justify intensified data sourcing in table 1.1. This table does not exhaust the calls for more data, but it does serve to illustrate types of purpose. I have organized it into four groups of data users: *researchers, clinicians, administrators/politicians*, and *industry*. Users do matter (Oudshoorn and Pinch 2003; Hyysalo, Jensen, and Oudshoorn 2016). Some individuals occupy several or all of these roles but may approach the data with different affordances over time, depending on their institutional setting and how it is governed (Pine and Bossen 2020). I propose the four groups as a way to reflect on how data are supposed to serve purposes associated with four types of good: *knowledge, health, good governance*, and *wealth*. These purposes weave into each other, just as individuals occupy multiple roles. Winthereik and colleagues have described how purposes often multiply when medical information is digitized: the digital form facilitates new ways of thinking about legitimate data usage (Winthereik, van der Ploeg, and Berg 2007). In that sense, we might say that purposes are not defined by human collectives alone. I return to the multiple purposes throughout this book because the simultaneous pursuit of several *different* goods—using the *same* data sources—is a key feature of intensified data sourcing. The most striking aspect of such a list is that so many people want to use data, practically all of which are produced, or will be produced, in the clinic or by patients.

What does this talk about "multiple uses" mean in practice? It means, for example, that when a hospital doctor sees a man with arthritis, a diagnostic code is used to describe the disease. In Denmark, the current computer systems use the international diagnostic manual, ICD-10, which is translated into a coding language. Codes are integrated into so many systems that updates to the international manual such as the ICD-11 (adopted by the World Health Assembly) are not immediately implemented (Green, Carusi, and Hoeyer 2022; Lie and Greene 2020). The diagnostic code specifies the type of arthritis in a taxonomy of disease. This code is combined with a code for the type of visit or health service and then used when communicating with other health professionals and when reporting to central registries.[2]

Table 1.1
Purposes for which actors want more data

| Actors/Goods | Stated Purposes Claimed to Justify Intensified Data Sourcing |
|---|---|
| **Researchers** Knowledge | Clinical research, such as the use of patient data to<br>• Document the effect of changes in clinical procedures and evaluate the effects (and side effects) of treatments<br>• Identify patients for randomized clinical drug trials or experiments with precision medicine<br>Health services research focusing on<br>• Usage of patient data to identify, e.g., patient needs or patterns of inequality<br>• Usage of patient data to assess, monitor, and compare, e.g., different forms of organization<br>Epidemiological research, such as the use of patient data to<br>• Monitor and predict disease burden and distribution<br>• Identify population health problems and potential causes of disease<br>Laboratory research, such as the use of patient data and samples to<br>• Validate new laboratory methods<br>• Carry out -omics research to understand mechanisms of disease<br>Big data research using population data sets, also beyond healthcare, to<br>• Identify patient pathways and interactions<br>• Identify patterns of disease or raise new hypotheses based on unexpected correlations |
| **Clinicians** Health | Systematic approaches to ensure consistency when dealing with, for example:<br>• Screening, triage and assessment of needs<br>• Follow-up (e.g., discontinuation of drugs, side effects, chronic care, polypharmacy)<br>Optimized treatment options<br>• Use of experiences from similar patients to make targeted treatment plans (e.g., in personalized medicine)<br>• Enhanced monitoring of patients at risk<br>Coordination and easy information exchange<br>• Communication and coordination between clinicians, across units and sectors<br>• Communication between clinicians and patients<br>Monitoring of quality, such as by using patient data to<br>• Identify safety problems (e.g., infection rates) and unusual patterns (e.g., in prescriptions or deaths)<br>• Document progress or achievement of goals<br>Public health surveillance, such as<br>• Pandemic surveillance<br>• Preventive measures |

(continued)

| Actors/Goods | Stated Purposes Claimed to Justify Intensified Data Sourcing |
|---|---|
| **Politicians and administrators** Good governance | *Political Level* |
| | Establishment of goals and assessment of whether goals are reached, such as<br>• Less use of physical force in psychiatry, or shortened waiting times<br>• Equality in health provisions and outcomes<br>• Medical error<br>Preparation for negotiations with external actors, such as<br>• Unions<br>• Pharmaceutical companies and other suppliers<br>Preparation of governmental reforms, such as<br>• Financing and remuneration models<br>• Creation of new units of administration<br>Prioritization between areas, such as<br>• Psychiatry or cancer<br>• Documentation of effect of services<br>Initiatives to cut expenditure and change routines, such as<br>• Optimization of resource use disruption of routines (e.g., using big data to predict, prevent, and plan)<br>Change of power balance<br>• Political management allowed access to the same knowledge as clinical staff<br>• Patient empowerment: patients are given access to their own data |
| | *Administrative Level* |
| | Remuneration and budget control<br>• Service per fee (e.g., general practitioners, pharmacies, and specialists)<br>• Value-based care<br>Accountability<br>• To avoid fraud<br>• To ensure quality<br>Efficiency and value for money, including "LEAN thinking" and measurement of effect |
| **Industry** Wealth | Create more efficient products, such as<br>• Personalizing treatment and prevention and achieving better effect<br>• Developing data tools that patients can use for monitoring and intervening in their disease<br>Expand innovation and lower research and development (R&D) cost, such as<br>• Use data as cheap "sandboxes"<br>• Minimize costs associated with RCTs<br>• Develop new products, including digital devices<br>Monitor effect of drugs postmarketing, such as to<br>• Identify side effects and drug interactions<br>Expand markets (and profits) with real-world evidence (RWE), such as by<br>• Moving indication thresholds or documenting off-label uses to enhance marketing<br>• Instigating value-based pricing through documentation of effect to differentiate prices<br>To use data as assets in their own right such as by<br>• Delivering data access and analysis to other companies |

The registries are used for research and planning. The code is the basis for reporting to quality databases used to document organizational performance. It is also used to produce lists of patients for organizational LEAN work. It is used for remuneration purposes, as well as to identify potential trial subjects for the pharmaceutical industry wishing to do randomized clinical trials (RCTs). This is what I meant when, in the introduction, I wrote that data have become *multiple* (Mol 2002): they are data on many things and come to produce (and form part of) very different phenomena than clinical care.

I now turn to examples from each of the four purposes that involve intensified data sourcing. I will show how data promises in relation to each of the four purposes are characterized by a peculiar temporality: the burden of evidence rests with the future, while "data and knowledge about what works" means relatively little in the present. Still, I also want readers to understand why people think they need more data. It is important to understand the challenges that researchers, clinicians, administrators, and industry representatives face to appreciate the drivers for intensified data sourcing.

## RESEARCH PURPOSES: DATA AS TOOLS OF KNOWLEDGE

Researchers are not one group with one set of interests, and the same data are used for very different types of research (cf. table 1.1). Still, across various fields, there is a push for data to gain statistical strength, capture complexity, and capture new aspects of phenomena or even establish new phenomena. There is also, for each research field, increased interest in accessing data for educational purposes. Data are crucial for career opportunities in an ever-more-competitive research world (Biagioli and Lippman 2020; Pinel 2021).

Medical researchers often work in the clinic, and they use patient data to identify patients for inclusion in trials or to document and publish the effects of local innovation. Health services researchers also depend on data generated in the clinic or by patients in order to explore the function and effect of the healthcare system, but often these researchers work elsewhere, in universities, semigovernmental, or independent research institutions. Similarly, epidemiologists must have access to population data to identify disease patterns and problems relating to, for example, inequality or discrimination. Epidemiologists also use data to identify agents of disease related to pollution or behavioral dispositions, among other topics, and to promote and document the effects of societal strategies of prevention. Patient data are also used for laboratory research, as

when new laboratory methods are validated, and in Denmark, clinically derived samples are also used for various forms of -omics research at the population level (Nordfalk and Ekstrøm 2018). On top of all these existing research forms, computer scientists now wish to explore amorphous clinical data sets with big data methodologies. They all need more data to generate future evidence.

Here, I wish to focus on just one instance of how investments in the creation of "future evidence" do not themselves necessarily involve using existing evidence in the present. My example revolves around the call for more data to enable what in the United States is typically called "precision medicine" and in Europe is known as "personalized treatment." The call for "personalization" is not new. Clark (1937) raised the problem of individual variation in drug responses as early as the late 1930s, and the shortcomings of RCTs being conducted primarily on white males have been addressed by policymakers and patient activists for decades (Epstein 2007). Nevertheless, this ambition has served as a lever for data integration projects throughout the Western Hemisphere from the 2010s onward, as laid out in the introduction. So how do researchers document their "need for more data"?

I have attended numerous scientific, industrial, and political conferences on personalized medicine. They are organized and attended by academic researchers, industry representatives, and healthcare system representatives. These actors also publish reports and strategy papers. The organization responsible for specialized care in Denmark, Danish Regions, drew up several such strategy papers, and based on one such "action plan" for personalized medicine (Danske Regioner 2015), a national project was launched (Sundheds-og Ældreministeriet and Danske Regioner 2016). This could be seen as a long process of consultations that leads to a synthesis—a perfect example of building on the best-available evidence. However, the project strategy endorsed by the health authorities provided few details on content, no work packages, and no methods (Skøtt, Rasmussen, Kruse et al. 2015). What the project strategy lacked in specificity was made up for in terms of the gravity with which the authors argue the urgency for investments in personalized medicine. One graphic in particular, which I will discuss in some detail, communicates this urgency.

The graphic is presented not only in reports, but also at conferences and in workshops, sometimes by several speakers during the same event—to the


FIGURE 1.1
The burning platform. A photo from one of the conferences displaying the graphic (reproduced on the left) that lists "Share of patients who gain no benefit from drugs" (referencing the FDA report). The graphic is used in reports and in an endless number of talks to argue the need for personalized medicine, here with the heading (in Danish): "The burning platform." (Credit for figure: Ulla Hilden. Photo: author)

point that it is introduced with variations of "you probably all know this figure," and often with the heading "The burning platform" (see figure 1.1).

The graphic lists "the share of patients who gain no benefit from drugs" in relation to nine diseases. It is said, for example, that 43 percent of Danish diabetes patients and 75 percent of Danish cancer patients gain no benefit from existing treatment options. It is an astonishing set of claims, especially considering that patients live with diabetes (and other diseases on the list) for many years. Why is nobody questioning these claims?

If lack of evidence for existing treatments is said to explain the urgency of investments in personalized medicine, a more pressing point, in my view, is that the graphic itself lacks 'evidence.' The Danish report from 2015, in which it first appears, references a US Food and Drug Administration (FDA) report as the source (US Food and Drug Administration 2013). The FDA report (and a very similar graphic) was used by President Barack Obama to promote personalized medicine in the United States. The FDA sounds like an authoritative source, but in support of the claims contained in the graphic, this report merely references an article from 2001 (Spear, Heath-Chiozzi, and Huff 2001). One perhaps could have expected more recent documentation on which to build the future of medicine. When

reading the 2001 article, however, the evidence base dwindles even more. The study presents no methodological explanation and actually just lists some numbers in a table in relation to an argument about what it would take to develop pharmacogenetic tests (see Senn 2015 for criticism from a statistician). Furthermore, there are some interesting leaps in each translation. The 2001 article referred to "response rates of patients to a major drug for a selected group of therapeutic areas," whereas the FDA report uses the caption "for whom drugs are ineffective." The leap from "response rates to *a* major drug" ("drug" as the singular) to the FDA phrasing "for whom drugs are ineffective" ("drugs," plural) to the Danish claim of "no benefit" is remarkable. In essence, this means that the claim that there is no evidence for the effect of existing treatments is itself being made without supporting evidence. Apparently, proponents of personalized medicine can make do with a narrative that merely *appears* to rest on data and a *promise* of better evidence in the future. The burden of evidence rests with the future; in the present, a good narrative will do.

The absence of interest in understanding the problem (the evidence base today) in order to promote a solution (personalized genomic medicine) involves a particular temporality. The solution is so grand that it becomes unnecessary to probe the problems supposedly haunting the present. And, really, why should scientists who are working in the nexus of science and policy want to debunk a graphic that might help raise billions in research funding for a large-scale data-gathering project? Despite all my reservations, I acknowledge that there are valid scientific arguments for embarking on large-scale projects aimed at a more data-intensive form of research. Researching clinicians rarely feel certain that a treatment will work: they sincerely want better predictions. I understand and appreciate their desire for data. My point is that the political process does not depend on "data and knowledge about what works" (as Deloitte suggested), and even the researchers involved in raising the funding do not seem to care. Researchers, clinicians, and policymakers coalesce in their pursuit of *future* sources of evidence while demonstrating tenuous relationships with *existing* evidence.

## CLINICAL PURPOSES: DATA AS TOOLS OF HEALTH

Like these researchers, most clinicians agree: They need more data! When you ask the clinical staff, at least in Denmark, they require data to ensure

consistency (e.g., in screening, triage, and treatment), optimize treatment options (precision medicine and monitoring techniques), ensure coordination (e.g., better documentation and communication), monitor quality (e.g., identify safety problems or document progress), and carry out public health surveillance (e.g., for prevention and pandemic preparedness). In short, they desire data and find them helpful when they facilitate action that optimizes patient care and prevention (cf. table 1.1). It is worth explaining some of these interests and the tensions they involve in just a little more depth.

Data in the form of code (e.g., diagnostic code) complement narrative documentation by delivering options for working systematically across patient populations, such as by identifying outliers in drug prescription patterns or medication errors. Codes are easy to search. They make it easier to schedule systematic checkups. GPs typically have around 1,600 patients, and they need tools to locate those in need of special attention. Similarly, in municipal care, hospital wards, and even screening programs, coded data help to build systematic routines. There are numerous examples of mistakes reflecting faulty routines or forgetfulness (Donaldson, Corrigan, and Kohn 2000). Anyone with responsibility for the life and death of patients is scared of making a mistake: data and digital support tools can help in establishing procedures to minimize this risk.

Another reason for clinicians desiring data in the broader sense of digital documentation is that should anything go wrong, they have documentation that can prove their innocence. Though Denmark does not have the same level of litigation as the United States, there is a driver for data intensification in what is called *defensive medicine*—data aimed at protecting the health professional rather than treating the patient (Wang 2017).[3] Clinicians are aware that data serve multiple purposes and cannot afford to think of them *only* as therapeutic tools. In clinical work, data are also legal tools.

Many clinicians also support infrastructural integration in the hope that it will ease their work with ensuring seamless patient pathways and coordination between units. As I will describe in more detail in chapter 3, they mostly loathe spending time on producing data. Very few health professionals got their education in order to make data for all the other users in table 1.1. Many clinicians, therefore, request better digital systems that can reduce the amount of time they must spend on data work. With a raised eyebrow, one doctor said to me during a break at a seminar about data tools in clinical practice: "It makes no sense that I can take this," he

pulled an iPhone out of his pocket, "and find a pizza restaurant in a second. But if I want to find any piece of information on one of my patients [in the electronic health record], I need to spend at least ten minutes!" Rikke, a civil servant working on the national integration of information technology (IT) systems, was painfully aware of these complaints and once explained to me how deep dives into some patient records revealed how some patients' heights had been recorded more than thirty times. Her guess was that it is simply quicker to ask the patients than to find the information in the record. Investments in digital data tools thrive on the promise of making the search easier, despite the common experience of being lost in a data overload that partly results from already having access to too much information. The sense of future potential overrules the current experience because the alternative is seen as accepting the status quo. Similarly, the prospect of being able to predict pandemics and plan public health interventions through massive data collection, including social media (SoMe) monitoring, survives repeated failures to make meaningful predictions and helpful interventions (Aiello, Renson, and Zivich 2020; Caduff 2015). Again, it might be because doing nothing is not a particularly appealing option when facing problems that affect the life and death of patients. Hope needs a nest in which to place its eggs, and currently data tools deliver such nests.

In some instances, the data produced while working with patients mingle in worrisome ways with governance needs. With Malene Bødker, I looked into a new data tool introduced in Danish municipal eldercare to document whether older citizens received the home care they needed (Hoeyer and Bødker 2020). For citizens requesting home care, the municipality had decided to offer rehabilitation instead. Rehabilitation meant that older people would be trained to be able to care for themselves. The municipality then installed a data-scoring system to measure whether these citizens, after this training, could manage well enough to get along without help. Rehabilitation was said to empower these older persons. The key promise was that the data tool would enable the staff to focus on patient goals and allow the older people to determine when they reached these goals. The municipality also paid for a scientific validation of the scoring system as a commitment to "evidence." However, the data were never used. The IT systems caused problems and could not deliver data in a format that the municipality could use for its analysis. Instead, the municipality decided to use data on *reduction in home-care costs* as proof of *rehabilitation*

*success*. The so-called success was backed with data, yes, but not data on older citizens getting the assistance that they need. In practice, older citizens were deprived of home care, and the data on their clinically measured ability to care for themselves—which were supposed to speak for them if they had unmet needs—remained silent in the databases. Despite the high-flying data promises, data meant to document clinical success and ensure accountability can be depressingly weak when confronted with strong political interests in reducing expenditures.

## POLITICAL AND ADMINISTRATIVE PURPOSES: DATA AS TOOLS OF GOVERNANCE

Clinicians and administrators subscribe to somewhat different data promises. I regularly receive invitations from high-ranking administrators and politicians wanting to discuss the potential for "total disruption" of existing practices. Clinicians typically have much more modest dreams, of smart software making their everyday workflow easier. Politicians and administrators, conversely, often talk about disrupting these workflows. In contemporary administrative parlance, "agility" refers to the ability to swiftly shift work goals, not agile and seamless tools that can underpin existing work practices and goals.

Not all administrators and politicians think about disruption. Some are just deeply concerned about finding ways for optimizing governance. Public authorities everywhere are tasked with providing a factual background on which political leaders can base their decisions, and with ensuring that political goals are translated into practice (Bessette 2001). Therefore, administrative and political data requirements cannot be separated. Good governance is both a political and administrative task. In fact, many of the political aspirations around data have been most clearly articulated in my interviews with administrators. The current vision for good governance involves the following objectives (also listed in table 1.1): to set goals and see whether you reach them; to be in a strong position when negotiating (e.g., with unions or industrial partners); to build reforms on "data and knowledge about what works"; to prioritize among areas based on analysis rather than the perceptions and opinions of clinical stakeholders; to use data to find ways of cutting expenditure and reorganizing care in smarter ways; and to change the power balance in the health services by giving patients

more leverage in relation to health professionals. Administrators are also inclined to think that financial control depends on a fine data grid. All of these political-administrative aspirations are in line with what is typically called "New Public Management (NPM)" (Hood 1991)—though by now, "Old Public Management" might be a more appropriate term. Politicians and administrators also consider it their job to build the infrastructures that fulfill the perceived needs of researchers, clinicians, and industry.

There is something undeniably appealing about most of these requests for data articulated by Danish policymakers (and also by politicians and administrators in many other countries), not least after four years where the White House was so ready to dismiss scientific evidence as fake news. In many ways, the essence of good governance is to emphasize accountability and efficiency. Some people might dislike authorities' interest in using data to prepare themselves for negotiations with trade unions, or for planning reforms aimed at cutting expenditures, but a public service that does not seek to spend taxpayer money with care is not particularly appealing either. Again and again, however, I have been struck by the disjunction between what data are supposed to do in the future and how they are being used in the present.

Data are said to be necessary to curb expenditures and provide better incentives. Is it through data that expenditures are currently being curbed—in the present? In some cases, perhaps, but certainly not as a general rule. For years, for example, a so-called reprioritization contribution was in place in Denmark, where operational expenses on hospital budgets were reduced by 2 percent based on an *expectation* of increased future efficiency (Højgaard 2017). This logic is not data-driven; it is a simple budget reduction. Arguably, it has made hospitals more efficient. Yet this was not based on data—it came from the necessity of running faster to care for patients with a smaller budget. Similarly, when building some new Danish hospitals, it was decided that the future buildings would be 8 percent more efficient than previous hospital buildings. This calculation was a key part of financing the investment, and yet the 8 percent reduction was not based on data or existing evidence (Ejbye-Ernst 2019a, 2019b). Similar "savings" were calculated into the budget when two administrative regions purchased a new electronic record system from the American supplier Epic (Allen 2019). This purchase is a story in its own right, and I will return to it in chapter 3, but the point here is just that the new system promised to deliver better data, and these data in

turn were promised to deliver savings—yet the promises themselves lacked any supporting data (Hildebrandt 2017d; Søgaard 2021). It turned out that there were no savings with the new system: on the contrary, efficiency went down, but the administrative management continued to insist on a future potential that just had not yet materialized (Wolf 2017). In short, there is no tradition for using "data and knowledge about what works" (again using Deloitte's promissory sentence) when implementing cuts in expenditure. It is a promise for the future, not a demand on the present.

Another way to create efficient use of resources is said to be by establishing the right incentives. Value-based care is one such international trend aimed at increasing efficiency (Hogle 2019; Bonde, Bossen, and Danholt 2018). It was already part of NPM to remunerate based on added value rather than stable budgets (Hood 1991). In that sense, the core elements of value-based care have a substantial history. In 1983, when President Ronald Reagan in the United States and Prime Minister Margaret Thatcher in the United Kingdom introduced neoliberal reforms, each in his and her own way, a group of Danish civil servants prepared a so-called modernization program. It was adopted by a right-wing government wishing to reform the welfare state (Nissen 2009). The program introduced market mechanisms to make the state more responsive to citizen needs and to make service delivery more efficient. It involved creating exchangeable units that could be delivered according to market principles, where each service could be measured and accounted for (Mau 2019). NPM was seen as the solution to unwieldly growth in the public sector, and already in the 1980s, proponents of NPM saw ICT as the key to developing more efficient and accountable management (Hood 1991). The emphasis on measurable units prepared the ground for datafication. Much later, and under the telling rubric "Forgive us, for we did not know what we were doing," the small group of civil servants behind the modernization program stated that they thought they were trimming a growing bureaucracy when in fact they were starting a bureaucratic mill of increased documentation (Gjørup and Hjortdal 2007). When value-based care is now presented as a new solution to the problems of NPM, it is a rhetorical trick more than any real change. Furthermore, the reinvention of value-based remuneration is not based on new "evidence." Policymakers refer to reports from the consultancies selling them software to make new calculations of "value," but these reports are themselves without citations

or anything that would count as empirical support by academic standards (SAS Institute 2017). The burden of evidence continues to rest with the future, not with the present when the investments are made.

While administrators continue to explain how data are needed to determine what is efficient, the cost efficiency of data collection is not itself considered. I asked a data analyst, Liselotte, about the cost of data sourcing, but she replied: "Basically, I couldn't care less about the cost [of collecting data]. . . . I just want [the data] that I need at my disposal." She wanted data to calculate the most efficient use of resources, but her own work did not need to be calculated on the same terms. I interviewed four civil servants from the national audit organization that works for the Danish parliament—the organization responsible for ensuring the efficient use of public resources. They all dismissed my question about the cost of data collection. One of them, Sarah, explained: "Data are not considered an expense. They are the basis for calculating efficiency."[4] In official reports and financial bills, cost items include digital tools, but rarely working hours. Clinicians just need to run faster when others want more data.

All the same, the current political conception is that accountable and efficient governance depends on having easy access to ever more data. This sense of data dependency is historically specific. Public healthcare has been governed from its inception—the question is how and by whom (Vallgårda 2003). In the early twentieth century, hospitals were in the hands of the medical profession (Vallgårda 1992). Doctors provided regular reports to the political system, but governance worked through delegation. Occasionally, doctors would request political guidance on how to make priorities (Wadmann and Pedersen 2020), but they would retain significant space for clinical judgment. Today, policymakers request Real-Time Data. Data help the administration overcome what has long been known as "the asymmetry problem," where only the clinical staff could claim to know what was happening on the ground (Kjærgaard, Knudsen, and Frølich 2008; Knudsen, Christensen, and Hansen 2008). Data thereby shift power balances and move decision-making power upward in the system, at least superficially (Moore 2018). Data promises exert political power, not because they build on "evidence" or always generate "evidence," but because they pledge a way out of the predicament—sometime in the future.

## ECONOMIC PURPOSES: DATA AS TOOLS OF WEALTH

All over Europe, policymakers increasingly see medical data as sources of *wealth* and economic growth, rather than just tools for health (Tarkkala, Helén, and Snell 2018). Therefore, they listen to the wishes of industry partners. As Sunder Rajan observes, states can be important agents for establishing novel markets in the area of health (Rajan 2003, 2006). In such markets, health data become a form of asset ready for capitalization (Birch, Cochrane, and Ward 2021; Vezyridis and Timmons, 2021). In Denmark, the medical industry is the largest export sector, and politicians facilitate and sometimes even pay for forums in which the lobbying can take place (Venkatraman, Mani, and Ussing 2015; Danske Regioner and Dansk Industri 2019). They also pay consultancies to investigate data needs in industry (PA Consulting 2019). In these reports, industry partners often raise concerns about "unnecessary" data protection, particularly challenges with the European General Data Protection Regulation (GDPR), which is seen as a barrier to "growth," and they complain about difficulties with finding the relevant data. Otherwise, the industrial representatives typically express the needs listed in the lower part of table 1.1.

As the industrial lobbyist Bent explained in the introduction, industry basically wants access to everything: easy, fast, and in a manageable and machine-readable format. Why? In particular, pharmaceutical and digital device companies express an interest in being able to "play around" with data to try out and develop new ideas (LIF 2020). They want access to clinical documentation and often talk about it as Real World Evidence (RWE) (US Food and Drug Administration, US Department of Health and Human Services, Center for Devices and Radiological Health, and Center for Biologics and Research 2017), a somewhat blurry concept with several meanings (Wadmann and Højgaard 2020). RWE is supposed to fuel commercial innovation processes. "Data sandboxes" is a common euphemism for this (Bech-Bruun 2019). "Sandbox" has a peculiarly innocent and childish ring, considering that it revolves around giving industry access to intimate patient data collected in the course of public healthcare that are now supposed to be used for the purpose of optimizing commercial market shares (sometimes the sandbox metaphor refers to synthetic or virtual data which are computer-manipulated datasets, though they still originate from patient data). It is not

any stranger, I guess, than using the term "cookie"—something sweet, pleasing, and gratifying—to refer to a digital surveillance tool (Bernsen 2019).

Besides "playing around," pharmaceutical companies also want access to patient data to identify side effects and drug interactions, and to expand their market by proving the value of drugs according to new pricing models associated with value-based care (where the impact on patient lives and health outcomes affects remuneration). The main pharmaceutical lobbying organization sees "a significant potential" in value-based pricing of drugs (Larsen, Hirsch, Broe, et al. 2020: 17). Some want to expand off-label use without conducting expensive RCTs, or they want cheaper alternatives to them (Eichler, Bloechl-Daum, Broich et al. 2018). Data access is also expected to help to magnify the market for drugs by finding indications for earlier diagnosis or expanded use (LIF 2020). Drug dispensing in chronic care is also increasingly dependent on data tools to monitor how patients metabolize and use drugs, and companies in areas such as diabetes care, therefore, invest heavily in digital health and machine learning to stay competitive (Karlsson 2021).

My interviews with industrial players have involved a series of rather special experiences. Few people in the private sector allow direct quotes, and I have rarely been allowed to record our conversations. In one case, I was first threatened with a lawsuit about a quote in a newspaper about surveys indicating public resistance to industrial access to patient data. However, in the end, we had several informative meetings, and I learned a lot. Another surprising instance was when a representative from one of the major global data analytics companies, a company sometimes referred to as a leading "health data broker," once invited me for coffee after I had asked a critical question at a public meeting. She wanted to know more about my work and discuss socially robust ways of using data. Having read about their notorious secrecy, I had long pondered how to approach this company—and here I was receiving an invitation from them! The commercial data world is full of surprises.

This mix of threat, openness, and secrecy says something about the ambivalent position of industry in relation to health data. It is both confident and defiant. It knows that it has the support of authorities and politicians, but also that it lacks similar support from the public. When I mentioned the term "data anxiety" in the course of a meeting (Cool 2019; Crawford 2014; Pink, Lanzeni, and Horst 2018), two commercial data analysts exclaimed with one voice: "Yes! That's it. We all suffer from data anxiety!" Partly, this

anxiety relates directly to uncertainties about how to comply with the rules of GDPR (an anxiety shared by many academic researchers), but my sense was that there was a deeper, more emotional aspect of their exclamation. They constantly had to tackle conflicting expectations to health data as both a source of revenue and as something beyond economic calculation (Hoeyer 2016).

As policymakers have responded to the sense of global competition, population data are turned into brands (Tupasela 2017b, 2021b), and brands subjected to optimization by way of removing "legal barriers for use" (Tupasela 2017a; Tupasela, Snell, and Tarkkala 2020). In 2003, Denmark introduced informed consent exemptions for biobank research and endorsed registry research without ethics approval to increase its research competitiveness (although in 2020, ethics approval was reintroduced for genomic data). In 1995, a registry was established where people could sign up if they did not want to be contacted by researchers. It was called the Researcher Protection Registry. In 2014, the ministry decided that too many people had signed up, and then the registry was simply deleted so all citizens again were available for research (Nordfalk and Hoeyer 2020). The minister responsible for this at the time, Margrethe Vestager, who later, as EU commissioner, became known for fighting the monopolies of the Big Tech companies, said that she had a clear feeling that people "had not meant to opt out of research" when signing up in the registry, and their research participation was necessary to create a basis for evidence-based policymaking (Nordfalk 2015). Since the justification of the deletion of the registry was not based on any data, it again appears that evidence is something that should materialize in the future rather than informing politics in the present. Legal entitlements in the present must be sacrificed to ensure future opportunities.

The competition among countries also is used by data brokers. In an interview with a global data analytics company, I was told that they always needed to see where they could do which type of study, in light of data availability. On their home page, I had noted that they seemed to possess any type of data imaginable, from social media to prescriptions, beautifully integrated in a graphic representation. In the interview, however, I learned that data sets are not easy to combine. Different countries provide different options for research, and no country has all the data included in their integrated marketing graphic. Social media data was on their menu, but they did not necessarily buy the data from the platforms. They could "scrape"

them ("we are pretty good at that," they said). Scraping has been described as a niche market in its own right (Angwin and Stecklow 2010).

Markets in health data take many forms, and these forms reflect both different infrastructures and the modes of calculating value that they facilitate (Franklin 2007; Callon and Muniesa 2005). Depending on the type, I was told, some datasets were available from the authorities, but not all types in all countries. The company expertise partly revolved around knowing where and how to get what. For example, they usually did studies depending on data from general practice in Sweden rather than in Denmark because GPs are publicly employed in Sweden. In Denmark, the majority of GPs are self-employed and not in the habit of pooling their patient data (at least not after a high-profile controversy about governmental reuse of data from general practice).[5] Sweden, conversely, is more restrictive than Denmark over access to registries and biobank samples.

Many government initiatives focus on attracting international investments (Ministry of Foreign Affairs, Ministry of Health, Ministry of Business and Growth et al. 2018; Schultz, Flyvbjerg, Thelborg et al. 2017). One government project costing 400 million DKK involved building an infrastructure for the identification of patients for clinical trials. It was meant to attract international pharmaceutical companies to carry out clinical trials on Danish patients. In an interview, the director of the initiative, Lise, explained the economic potential to me, but when I asked for data or calculations that indicated a positive return, she retracted and said: "It's actually not like we earn money on this," and then continued, "it's more about learning from industry and so on." In fact, she now wanted clinics to collect more data to document the learning value of trials. Again, the promise of economic growth does not need to be supported by "data or knowledge about what works" to ensure investment. Data are expected to deliver future evidence for the decisions already made, not to inform decisions in the present. The recruitment initiative was made permanent with additional funding. Similarly, the national audit organization did a survey of investments in digital infrastructure and found that in 23 out of the 96 projects they investigated, the risk calculation was produced *after* the investment was made (Statsrevisorerne and Rigsrevisionen 2020). Data are often produced to support what is already decided, not to inform decisions.

Some of the big American tech companies also seem to have an interest in the Nordic countries. They place data centers in Denmark, for example, in

order to "green wash" their electricity drain (Maguire and Winthereik 2021). In an interview with the investigative journalist Markus Bernsen, Rob Nail from Singularity University further explained that the reason why the US big tech companies invest so heavily in Denmark is because this tiny country is considered an interesting "testbed" for their vision of "a fully automated society" (Bernsen 2019: 81). Denmark is far more digitized than most places, he said, and it also has a social security system that can ensure stability while existing societal structures are "disrupted." The companies aiming for total disruption expect some form of upheaval when people lose their jobs, and when their options become based on automated predictions in areas such as education, health, and policing. Here—as the logic goes—the Danish welfare state might stabilize the conditions for the experiment.

In the area of health, access to Nordic data resources is also interesting for very concrete reasons. These resources can be used to validate data prediction that otherwise is based on commercially available data, such as those from Facebook (or Meta), Apple's Research Kit, and FitBit. Teis, a centrally placed IT developer, knew firsthand about the interest in Danish health data among several US big tech companies. The companies, I was told, want to validate the predictions they make based on data generated on their own platforms. The health data of Danish citizen-patients are authenticated by health professionals using the same traceable personal identity numbers as Danish phone users provide to the phone companies. This type of information can say a lot about the validity of the predictions that otherwise are just based on inferences from consumer-derived data. Coming from the opposite direction, as it were, Danish researchers and clinicians dream of access to the data on people's phones and their online lives, also in order to integrate the two sources of health data.

Big tech companies and pervasive state tracking sound, to some, like a match made in heaven. The platform offering online access to health records, *sundhed.dk*, has a strategy to facilitate citizen upload of consumer data such as that from wearables. It is said that the purpose is to ensure citizens' safe storage. It also facilitates the linkage of data. The policy papers on a new European Health Data Space also promise the option of integrating consumer-generated data with public healthcare data, as mentioned in the introduction. Not everybody is content with the speed of the public initiatives, however, and several projects (e.g., something called the HedaX project) have received large amounts of public and private research funds to build

infrastructures where the personal identifier number can link data on shopping from loyalty cards with activity-tracking data (e.g., FitBit), and commercially purchased tests (e.g., ancestry testing), as well as public health records. If commercially available data can be linked with hospital data, companies can claim higher accuracy of the predictions they sell—also in other countries and for persons who have never agreed to merge their FitBit data (for example) with medical records.

Finally, there are markets for health data that operate underneath the surface (Tanner 2017). Some people working with data security at the national level have told me about the black market for health records created through hacking. Apparently, medical records sell for as much as 1,000 US dollars, while credit card information can be purchased for just 1 dollar. The buyers are, ostensibly, mostly people using the information for blackmail, but insurance companies also are said to be among the customers. They want access to unfiltered health records to test the veracity of hospital invoices and patient claims (see also Rowley 2017). Unfortunately, their attempts to protect their legitimate businesses by detecting patient and hospital fraud help generate a black market for data, thereby putting other patients at risk.

## PARADOXICAL PROMISES: THE EVIDENCE-FREE QUEST FOR FUTURE EVIDENCE

In this chapter, I have explored how the data promises promoted by policymakers, consultancy companies, and data enthusiasts at all organizational levels interact with perceptions of data needs among researchers, clinicians, politicians, administrators, and industry representatives. I have shown how the promise of *future* accountability based on data is a key driver for intensified data sourcing, but also how data do not deliver such accountability in the present.

Although Deloitte might proclaim an ambition of having political and organizational decisions based on "data and knowledge about what works," neoinstitutional organization theorists will describe such as vision as being slightly naive (Weiss 1986; Brunsson 1989; March and Olsen 1976). This phrase from the Deloitte report probably should be read more like a marketing slogan than a sincere ambition. A lot of organization studies have documented how organizations tend to mimic decisions from other organizations rather than conducting their own data analysis (Abrahamson 1996; Benders and van Veen

2001; Czarniawska 2005; Frenkel 2005; Kieser 1997; Newell, Robertson, and Swan 2001). It is not even obvious that organizations, or governments for that matter, *could* base decisions on data. Most of the decisions about data investments mentioned in this chapter touch upon volatile parameters and affect stakeholders with competing values. Data can inform decisions, but decisions cannot be data-driven (Rittel and Webber 1973). As the organization theorist Charles Lindblom once said, most organizational decisions build on nothing more than the "science of muddling through" (Lindblom 1959).

Feldman and March (1981) have argued that organizations tend to collect—and request—more data than they use because this enhances the legitimacy of those in power. It is easier to accuse top management of having collected too little than too much information. Alvesson and Spicer (2012) even claim that most organizations thrive on collecting data without using them. It creates what they term "functional stupidity":

> Functional stupidity is organizationally-supported lack of reflexivity, substantive reasoning, and justification. It entails a refusal to use intellectual resources outside a narrow and "safe" terrain. It can provide a sense of certainty that allows organizations to function smoothly. (1196)

Alvesson and Spicer (2012) are very clear that functional stupidity does not operate through "intellectual deficits but through political expediency" (1214). Against this background, and in light of the examples across all four intertwined domains, it is time to rethink the work of data promises.

Data investments thrive on a claimed need for future evidence, but investment proposals are rarely backed by similar evidence themselves. Data are not used in the present, but they are expected to do their trick in the future. It is a paradox where both stories are true: people like to think of their decisions as building on data and "evidence," at least in the future, and yet they also neglect using data or "evidence" in the present. I suggest that the coexistence of two truths can be productive for policymakers because it provides an aura of legitimacy and allows *postponement of action*. By claiming that it is better to wait until more data have been accumulated, data promises generate a form of temporal disruption of public accountability. In this context, data-as-promise is—politically—a more powerful resource than data-as-evidence. The paradox also is productive for researchers, who can get funding for research by playing along with popular narratives and still adhering to their scientific values, if not in the present, then—perhaps—in the future. It is productive for clinicians, who hope to get tools in the future that they are

missing in the present, and who can adhere to values central to evidence-based medicine without being able to reach that threshold here and now. Further, it is productive to industry, which can use it as leverage to get access to resources.

It is necessary to take both tenets of the paradox seriously and not see the desire for data as somehow fake or akin to lying. People have good reasons for wanting data even when they do not use them for making decisions to invest in data sourcing. Their "data needs" do not arise out of the blue—they are nested in real-world problems. It is, for example, no coincidence that the tenets of New Public Management (NPM) can be reinvented again and again (Pedersen 2019a): the problems of accountable and efficient government are not solved. While the sense of data needs becomes a motor for data collection, the act of postponement makes it possible to *feel* accountable—in the present—although it does also push the fulfilment of accountability ideals into the future. Promissory data legitimize postponement of action—and *fill the waiting time* with a genuine and serious task: intensified data sourcing.

All in all, the eagerness of policymakers to attract capital and make themselves available as testbeds for digital disruption reflects the rhetoric of economic threat from the Boston Consulting Group report cited above. Consultancy reports, policy papers, and think tank statements all argue the need for ever-more-radical measures and lenient legal frameworks, "lest we lose competitiveness" (Regeringen 2019: 7). If the anthropologist Veena Das once noted that citizens in low-income countries are often forced to create an opportunity out of economic deprivation—for example, by making their bodies available for organ harvest (Das 2000)—then citizens in the Nordic countries are increasingly forced to *create vulnerability out of privilege*, "lest they lose competitiveness." It is not their bodies but their data that are up for sale, and they are not making the bargain themselves. Their governments are managing the business for them. Vezyridis and Timmons (2021) have made a similar argument about health data in the United Kingdom. Petersen (2019) suggests that the new trade in data might be a "Faustian bargain" of surrendering something as valuable as one's soul to reach out for the promised land of digital health. My point here is that the nature of the bargain is unknown because the question of which goods the stakeholders will use data to pursue remains unsettled.

I have focused on the data needs articulated in reports and by the people I interview, but I could also have included their negative counterparts. When I mention researchers wanting data to deliver new knowledge, some say that researchers just want data to boost their curriculum vitae (CV) in an academic world corrupted by metrics. When I mention clinicians wanting data to deliver health, some say that many clinicians just want their clinic to have "good quality" in malfunctioning systems for performance measurement. When I mention administrators and politicians wanting data to deliver good governance, some say that they are really trying to build structures for surveillance and control, akin to what you find in cruel dictatorships. When I mention industrial actors wanting to generate wealth for society, some say that it is an economy driven by a form of greed that, like a cancer, eats its way through the bone and marrow of medical evidence.

I have not described these negative counterparts in any detail because they are motives that people attribute to others, never to themselves. Some also point to police uses of medical data. Still, the police can access genetic data in Denmark only for investigations of terrorism cases, not for other types of crime, and it has never appeared as an official reason for any of the initiatives. The negative counterparts, these less benign motivations, are not the official reasons for wanting data. Still, once data are available, they propel new uses, and this sense of "potentiality" (Taussig, Hoeyer, and Helmreich 2013; Svendsen 2011) might very well be a driver for intensified data sourcing. In that sense, the negative counterparts can also be understood as what my interviewees see as likely implications of intensified data sourcing, although it is what they fear rather than what they desire.

How do the data promises discussed in this chapter interact with local infrastructures and affect the everyday lives of patients? This is the topic of chapter 2, which also explains the Danish health data infrastructures in more depth, and thereby also their appeal to the big tech companies and to policymakers elsewhere. By moving closer to the infrastructures shaping everyday living in Denmark, I provide a better sense of the context for the developments I describe in this book. Context is also essential for anybody who wants to understand why policymakers from all over the world, who dream about an integrated health data infrastructure of the Danish type, might not get exactly what they expect when investing in digital infrastructures.

This is a section of doi:10.7551/mitpress/14926.001.0001

# Data Paradoxes

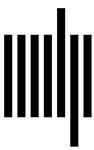
## The Politics of Intensified Data Sourcing in Contemporary Healthcare

By: Klaus Hoeyer

## Citation:

*Data Paradoxes: The Politics of Intensified Data Sourcing in Contemporary Healthcare*
**By:** Klaus Hoeyer
**DOI:** 10.7551/mitpress/14926.001.0001
**ISBN (electronic):** 9780262374156
**Publisher:** The MIT Press
**Published:** 2023

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press


The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Hoeyer, Klaus, author.
Title: Data paradoxes : the politics of intensified data sourcing in contemporary healthcare / Klaus Hoeyer.
Description: Cambridge, Massachusetts : The MIT Press, [2023] | Series: Infrastructures series | Includes bibliographical references and index.
Identifiers: LCCN 2022019995 (print) | LCCN 2022019996 (ebook) | ISBN 9780262545419 (paperback) | ISBN 9780262374163 (epub) | ISBN 9780262374156 (pdf)
Subjects: LCSH: Medical informatics—Denmark—Case studies. | Medicine—Denmark—Data processing—Case studies. | Medical policy—Denmark—Case studies.
Classification: LCC R858.A3 H64 2023 (print) | LCC R858.A3 (ebook) | DDC 610.28509489—dc23/eng/20220830
LC record available at https://lccn.loc.gov/2022019995
LC ebook record available at https://lccn.loc.gov/2022019996