Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-020-00535-6 OPEN

Data promiscuity: how the public–private distinction shaped digital data infrastructures and notions of privacy

Klaus Hoeyer [1]✉

This essay discusses the performative effects of the public–private distinction on digital data infrastructures in healthcare. The words 'public' and 'private' hold many meanings. This analysis focuses on how they are used both in an informational sense (what is kept secret or strictly controlled versus what is out in the open or shared) and an institutional sense (issues of ownership and purpose such as being state-owned and governed for the common good or privately owned and aimed at generating profit). In the political construction of digital infrastructures, the two senses are deeply intertwined: changes in relation to ownership and purpose affect what is kept secret and what is shared. Furthermore, when policymakers search for ways to protect one aspect (privacy) they sometimes opt for tools from the other (by conceiving of data as private property). The informational interconnectedness facilitated by digital infrastructures produces a form of 'data promiscuity'. Data promiscuity is a condition where data are indiscriminate in the choice of partners: what is seen as data on a thing for one purpose can always become seen as data on another aspect of that thing and be used for another purpose and by another user. Data are set free to pursue gain or pleasure, but this freedom involves certain dangers for the persons from whom they derive. Data promiscuity is the contemporary condition of possibility for health research. By unpacking the wholesale categories of public and private through which the contemporary situation came about, there is a better chance of rethinking the problems it involves, and for suggesting new solutions to ensure social sustainability. The argument is based on developments in one of the most fiercely digitalised and datafied countries in the world: Denmark.

[1] Department of Public Health, University of Copenhagen, Øster Farimagsgade 5A, Room 10.0.09, 1014 Copenhagen, Denmark. ✉email: klho@sund.ku.dk

Introduction

In the course of just a few decades, pervasive digitalisation has facilitated an unprecedented datafication of many aspects of modern life (Ruckenstein and Schüll, 2017; van Dijck, 2014). This datafication has created new pathways of information exchange and involved transformations of economic and political infrastructures. With this paper, I suggest that these changes have been influenced by the work of a *public–private* distinction that social scientists and political actors use to describe societal problems, as well as to invent political solutions. The public–private distinction, however, not only describes problems and solutions; it *prescribes* them by way of shaping policy thinking. The distinction provides, in Geertz' (1973) sense, both a model *of* and a model *for* reality. The distinction is nonetheless engrained in a cultural ideology that poorly captures the multiple and complex interests of the citizens subject to datafication. It is therefore time to look at what it has produced as a model *for* infrastructures of digital datafication and to pursue new solutions.

While the public–private distinction is used in many senses,[1] two of them stand out in relation to digital data sharing (Ariès, 1989; Johansen and Andrews, 2016); one relating to *informational* aspects (what is kept secret or strictly controlled versus what is out in the open or shared) and the other to *institutional* issues of ownership and purpose (typically seen as a matter of being state-owned and governed for the common good, or privately owned and aimed at generating profit). These two aspects are particularly interesting to look at in tandem because a datafied economy implies transformations in both the informational (who gets access to what) and institutional (who owns, operates and profits) aspects of emerging data infrastructures. This is perhaps most forcefully seen in the EU General Data Protection Regulation (GDPR), which is explicitly aimed at a double agenda of promoting a single market for data (institutional transformations) while enhancing privacy rights (informational transformations).

Today, 'data' on people have become embedded in open-ended digital networks. In such networks, data are always subject to potential viewing. They are 'promiscuous' in the sense of being relatively *indiscriminate* in their relations to users: what is seen as data on a thing for one purpose can always become seen as data on another aspect of that thing and be used for another purpose and by another user. Data are relational: they are not data on any one thing, but on whatever a given analysis uses them to investigate (Leonelli, 2016). Accordingly, there is no clear definition of *health* data anymore: any type of data can become data on health when used to predict a health outcome. Big data analytics has undermined any clear subject matter delineations. Furthermore, Prainsack (2019) points to the multiplicity of data, as being 'able to be in more places than one at the same time, in leaving traces even when they are 'deleted', and of being able to be copied and used by several people at the same time, independent of what the others are doing' (p. 13). This multiplicity gives rise to promiscuity.

Promiscuity is typically associated with indiscriminate choice of sexual partners and though this practice is admittedly associated with certain elements of pleasure, the risks of unintended offspring and dangers to health has made the promiscuous person subject to moral condemnation. In response to public concerns about data privacy, the GDPR and related forms of policy work similarly tend to make individuals responsible for restraining and controlling data relations. As a result, the exposed data subjects can be blamed for lack of prudence and self-restraint. Data monogamy is presented as the safe choice for individual citizens, just as self-restraint has been the dominant public advice guiding morality on sexual desires throughout history. Data infrastructures, nonetheless, incentivise data sharing—or incentivise data promiscuity, as it were. Therefore, there is a need for tools other than encouragement of individual (self-) control to protect citizens in data-intensive societies. Data monogamy is hardly a choice, when 'data sharing' is simultaneously becoming a precondition for societal engagement.

To develop better tools, I suggest unpacking the broad categories public and private to see how they shape data infrastructures. A good place to start investigating the performative effects of the public–private distinction on digital datafication is Denmark—a small European welfare state, which is among the most radically digitized and datafied countries in the world (Hoeyer, 2016). Denmark is a small welfare state in northern Europe with just 5.7 million inhabitants, but it is a country taking the role of dedicated digital frontrunner. Public registers, electronic health records, social services and commercial services such as banking, telecommunication and transport options all operate in digital information infrastructures using the same traceable identity numbers. For decades, the Danish health data infrastructures have received international praise as essential resources for epidemiological research (Frank, 2000, 2003). Today, they are seen also as resources for economic growth (Hoeyer, 2019; Tarkkala et al., 2019). To understand the conditions of possibility for health research in a digital age, it is therefore time to understand the wider transformations of data infrastructures through digitalisation.

After a brief discussion of the public–private distinction and a note on methods, I outline first how the public–private distinction has informed a political ideology that has changed Danish welfare state infrastructures, including the health services. A programme of New Public Management has interacted with fierce digitalisation and a very data-intensive form of government and clinical care. In the next two sections, I turn to the implications of these transformations: first with respect to the informational aspects of the public–private distinction, and then the institutional ones. In this way, I explore not only the performative effects of the distinction, but also the concerns and interests that it is typically used to denote. In conclusion, I discuss the need for moving beyond the solutions that public–private thinking tend to offer to address those concerns.

The public–private distinction

In the course of considering the enduring interest in negotiations of secrecy in light of changing institutional structures, Jeff Weintraub suggests that the public–private distinction can 'neither be conveniently simplified, nor usefully avoided' (1997:38). The point is to become aware of how it shapes political logics, and to shift the focus from what the public–private distinction *means* to what it *does* in our contemporary data economies. A first step in this direction is to acknowledge that public and private are not ontologically separate domains. In a seminal article on the public–private distinction, feminist linguist Susan Gal once argued that though it is common to talk of a 'blurring of boundaries' between public and private, such talk relies on a basic misunderstanding (Gal, 2002). Public–private thinking operates as a *'fractal distinction'*, she argues, meaning that any segment that has been divided into either public or private can itself be subdivided again into its 'public' and 'private' aspects. As other fractal distinctions (e.g., nature/culture, female/male), the public–private divide is a sense-making device shaping the world that it is said to classify: public and private are performative concepts that align activities with meanings and thereby shape people's behaviour.

Gal specifically suggests exploring how ideas about public/private can become institutionally engrained and create moral zones with special rules. She provides an example from the era of

communist Eastern Europe where taking something from what was considered the public realm could be experienced as a smart appropriation of resources, while taking something in a similar way from a realm deemed private was considered theft. The public–private distinction in this way defines moral responsibilities, and it is politically effective because it is perceived as merely descriptive and referring to ontologically distinct categories.

The institutional transformations that I explore in this paper revolve around data infrastructures and therefore they have implications not only for ownership and moral attributions of rights to profit, but also for the informational aspects of the public–private distinction; for privacy. This takes us to the informational aspects of the distinction. One of the first, and certainly one of the most influential, legal treatments of a 'right to privacy' is an article by Warren and Brandies (1890). It came as a response to the invention of photography and its use in gossip magazines to convey unsolicited images of people. Ideas about a right to be left alone in 'public' spaces have a history preceding the camera, though. In his classic 1974 book *The Fall of Public Man*, sociologist Richard Sennett argued that big cities gave rise to a feeling of a right to pass others on the street while pretending not to be noticed. It was the unavoidable closeness of the city that generated a pretence of distance. Today, digital platforms and health data infrastructures similarly create unavoidable closeness as people cannot avoid leaving data traces that can be seen and used by others.

What counts as a privacy intrusion is not stable, and ideas about privacy have undergone drastic transformations throughout history (Ariès, 1989). While people have always cared about what they share with whom, ideas about how to make that choice differ significantly over time. As sociologist Georg Simmel (1950a) asserted, the basic dialectic between what is kept secret and what is manifest shapes every society. He also argued that secrecy is never ultimate. Instead, secrecy is about the means available to people to control what is shared with whom. Moreover, indeed, digital infrastructures influence that political struggle in ways Simmel could hardly have foreseen.

To explore the work done by the 'public–private' distinction, I now turn to the digital infrastructures in Denmark. I was inspired to write this paper in the course of ongoing fieldwork where I combine policy analysis with participation in meetings, workshops and conferences, as well as interviews with stakeholders in ministries, health data organisations, and associations for industrial partners. I also interview people involved in setting up data infrastructures, data analysts and activists engaged in data-intensive healthcare practices. I have participated in more than 35 events and conducted >60 interviews, and along with these activities I have gathered a lot of material (reports, news stories, policy papers) on ongoing digital transformations. For this paper, I have traced the history of the involved institutions discussed by the informants building an archive of material from homepages, minutes, news stories and interviews. It is impossible to deliver a comprehensive overview of all of the transformations and I can only present some examples from the wider archive to illustrate what I have come to see as general trends. I first illustrate how the public–private distinction together with digitalised datafication have changed Danish welfare state infrastructures.

'Privatising' infrastructures: Denmark as digital frontrunner

The modern welfare state came into being in tandem with the construction of infrastructures for communication, energy supply, and transport; as well as strong informational infrastructures to ensure adequate statistics indicative of population needs (Sætnan et al., 2011). In Denmark, as in the other Nordic countries, the infrastructural agencies were originally owned by the state or municipalities. Beginning in the 1980s, the agencies supplying many of the infrastructures have turned into state-sanctioned, but mostly commercially-run companies. The governance of the services has shifted from a wider social contract between state and citizen to contractual relationships outlining the distribution of tasks, rights and duties between citizens (as users), companies (as service providers) and the authorities (as custodians).

In Denmark, this transformation in ownership and contractual form was set in motion with the Danish version of New Public Management, the so-called 'modernisation programme' from 1983 (Finansministeriet, 1983). This programme sought to model the 'public sector' on ideas associated with the 'private sector', in order to achieve benefits associated with 'private' companies, namely efficiency through competition. The 'modernisation programme' corresponds in many ways to what Thatcher initiated in the UK and Reagan in the US, and its implications have been scrutinised by governmentality studies and related forms of social critique. It created a new demand for data to make 'public' service delivery respond to performance measurements in pursuit of 'market' mechanisms. As elsewhere, changing governments also gradually privatised the ownership of various infrastructural agencies. The national postal services (Post Danmark, later PostNord), telecommunication and internet delivery (KTAS, later TDC), energy supply (DONG), and the operation of public transport options have, one by one, been turned into companies.

If the modernisation programme created a data-intensive form of government, it could do so by drawing upon already existing data infrastructures. In 1968, Danish authorities for taxation purposes established a central identity register, the Central Person Register (CPR), which assigned each citizen a ten-digit number. All property and all companies were similarly given a number. Once in place, the CPR number has come to be used in all encounters with public services, as well as increasingly in the commercial sector. The ability to validate the identity of citizens with high certainty, has made the CPR an attractive partner for a constant flow of new users and for new purposes way beyond taxation. Today, the CPR number is offered for use according to a payment model for commercial services in need of user identity validation (achieving the identity verification that companies like Facebook and Google aim for). The CPR systems in the Nordic countries have facilitated the generation of some of the most comprehensive databases in the world with an ability to combine educational, occupational, income, taxation and health data in long time series (OECD, 2013).[2] The CPR is an essential resource for health research allowing investigation of the interplay of social, economic, environmental and medical factors. Most Danes probably consider the CPR system 'public' and the whole-data-intensive form of government as indicative of the goals of the welfare state. The homepages describing the register would give no other impression. The running of the register, however, has for years been outsourced to an international IT company, CSC,[3] and private sector actors that use the register are not required to make their data available for public research.

As the Danish form of government is so data-intensive and tracks each citizen in great detail, it has created a push for digital technologies to handle all the data (see also Andrejevic, 2013). Digitalisation is and has for decades been a high political priority in Denmark and is also acknowledged as essential for ensuring the availability of the data needed for New Public Management (Forskningsministeriet, 1994). State ambitions of ensuring Denmark a position as digital frontrunner are not limited to the so-called 'public' sector. Already in the 1960s, the Nordic banks collaborated with public authorities on the digitalisation of the financial sector. A computerised billing method known as PBS

(da; Pengeinstitutternes Betalingsservice) was introduced in 1968 in Denmark, and in updated formats it today ensures an almost fully digitalised billing system. In 1983, a national credit card (Dankort) introduced digital payment and Denmark quickly had the highest number of digital payments in the world. In 2000, a national digital postal service system called E-Boks was invented. In 2005, state authorities decided to use E-Boks for correspondence to all public employees and in 2012, a law made it the default means of communication with all citizens. It is also used in the health services and for invitations to participate in research. In 2003, a national secure access and identity system called NemID was developed by a consortium of companies in response to a public tender. Today, it is used by both authorities and commercial service providers. E-Boks similarly came about through a government initiative but operates as a shareholder company partly owned by another shareholder company, NETS, which is also behind the digital payment option, the Dankort and NemID. In fact, most of the 'public' digital services are owned by or outsourced to various combinations of a select group of companies, including, NETS, IBM, CSC and KMD (all known under their abbreviations).[4] In consequence, Danish citizens today consistently interact with what is described as 'public' services (authorities) through 'privately' (i.e., commercially) delivered digital interfaces.

In sum, digitalisation and datafication have emerged through not just technological developments but also as a consequence of the urge to modernise the state based on ideals associated with the 'private' sector. As digital infrastructures for 'public data' are outsourced to, run by or maintained by commercial companies, it makes little sense to see public and private as ontologically distinct domains. Rather, notions of public and private legitimise certain forms of organisation. Public/private thinking justifies that the infrastructures are run commercially to achieve values associated with the 'private sphere' such as 'innovation', 'efficiency' and 'digital competence' (see also Mazzucato, 2015).

After having in this way outlined the infrastructural transformations enacted in the image of the public–private distinction and in pursuit of digitalisation, I now turn to the implications of these transformations by looking at the aspects of life that the public–private distinction is typically used to describe. I begin with the informational aspects (what is kept secret and what is manifest and to whom) and then in the subsequent section explore institutional aspects (ownership and attribution of praise and blame).

Informational implications: the secret and the manifest

The digitalisation of data infrastructures intervenes in intriguing ways in the dialectic between the informational aspects of private and public, or what Simmel (1950a) called the secret and the manifest. Digital data integration and the more complex networks of commercial suppliers open up access for many more potential users than the old closed filing cabinet did in the offices of authorised civil servants or health professionals. The paper-based medical record used by the general practitioner in the old days primarily served as a local memory-assisting device. A digital database, in contrast, always has multiple users and an in-built potentiality (Winthereik et al., 2007). There have always been ways of peeping into files; what the transformations above have done is to change the conditions under which people negotiate the secret and the manifest and the complexity they need to manoeuvre to do so.

Data infrastructures are designed to both facilitate and control data flows (Andrejevic, 2013; Ruppert et al., 2017). The Danish authorities are well aware that data integration aimed at giving new actors access also comes with a risk of unauthorised viewing or even data leakage. Accordingly, they typically emphasise 'data security' as an important feature of all policies (Danske Regioner, 2017; Lindstrøm et al., 2017). This sounds simple, but it is not. On the one hand, total data security is an illusion; on the other hand, what authorities deem 'safe' in the sense that only authorised personal access data, might still involve making certain pieces of information available to selected staff members that individual citizens would have preferred keeping secret. Data integration is about changing who may know what about whom, and technical data security cannot resolve this type of basic political disagreement (Wadmann and Hoeyer, 2018).

The intricate implications for informational privacy of digital integration of health data can be illustrated with Denmark's e-health portal (Sundhed.dk). This portal integrates access to medical records, lab results, pharmacy prescriptions and other forms of data. In some surveys it is rated as one of the most advanced in the world (Frost and Sullivan, 2017). In the name of empowerment, citizens have also been given online access to their health data through the portal: the policy logic was that data 'belong' to the citizens (note the metaphor of private ownership), and therefore they should have access to them. Interestingly, the e-health portal had to invent a so-called 'privacy' function to allow citizens also to limit which elements that can be viewed online. The privacy function allows citizens to hide material not just from health professionals, but also—and more surprisingly—from the citizens *themselves*. Data remain available for administrative and research purposes. The function does not ensure total data monogamy.

Why would people keep data hidden from themselves? It turned out that some patients felt forced by family members, employers and insurance companies to share a printout of their data now that their whole medical history was so easily available to them. Hence, this form of 'privacy marking' in effect has come to mean keeping data *away* from the individual whose privacy is a stake, rather than in the hands of this same individual citizen. 'Ownership' of the data was not empowering, when it meant others could exert powers over the owner. No access was a better form of protection. To keep data 'secret' is in this case to keep it in the 'public' database only, out of reach from the data subject.

Of course, Danish digital infrastructures also leak in the classic sense. In fact, there is a continuous stream of media stories about data leaks. The most notorious leaks are those about famous people having their secrets exposed.[5] One such leak gained enormous press coverage following the revelation that a tabloid magazine called *Se og Hør* illegally bought credit card information on famous people. In English, the magazine's name would mean 'See and Hear'; and it is indeed a magazine specialising in privacy intrusion in the old sense of Warren and Brandeis (1890). The magazine's snitch was a man employed by IBM to do work for NETS on the maintenance of the national credit card, Dankort, and the access code system, NemID. Through him journalists could trace, for example, the holiday destinations of the royal family or where celebrities were giving birth (Retten i Glostrup, 2016). The digitalisation of the financial infrastructure thus gave multiple actors access to information, including health information: a state-sanctioned privately run infrastructure gave multiple points of access to intimate data and made them subject to (illegal) trade.[6] The infrastructural transformations had informational implications.

When I have interviewed patients, they are rarely concerned about data leaks of the type broadcasted with the *Se og Hør* revelation. They are worried only when they are confronted with their own data in unexpected places or used against their own interests, as when a doctor mentions having noticed a psychiatric diagnosis in their record while actually treating something else. In an interview I conducted together with (colleague's name deleted

for anonymity), a woman with a chronic disease thus talked about sometimes feeling uncomfortable when doctors or civil servants mentioned things from her files that she herself found irrelevant for the specific condition she wanted to discuss. She then said something that has stayed with me. Even if it might be unusual, I believe it illustrates an interesting point about the kind of work that the public–private distinction performs as people seek to separate the secret and the manifest in the new digitally integrated data landscape:

> I have a feeling that the day I got diagnosed, many years ago, I became public property. I have some kind of obligation towards my society; there is some kind of *right* to step in and pry into my life. It is a combination of all sorts of registers where you can extract all sorts of information about me that others are not obliged to deliver about themselves, as persons.

Although she says 'I have some kind of obligation', she simultaneously questions this obligation. She feels infringed, and relates it to 'prying'. Hence, she gives the impression of moving in a landscape of obligations that she does not acknowledge as legitimate. One of the things that makes this quote remarkable is that she worked for a patient organisation to promote data integration and to improve collaborations between data-responsible authorities and the pharmaceutical industry. She articulated the 'private' realm as one of clever innovation where she felt patient organisations could engage in actual negotiations of research content. So, the experience of being 'public property' was associated with a sense of duty and a passive patient role marked by lack of control, while her sense of 'private company' related to experiences with actual involvement. 'Public' use of data was prying, while 'private' was sharing. Again, public and private are performative concepts that intervene in the dialectic of the secret and the manifest. They can make people provide data out of duty in the 'public sector' and then promote further data sharing with the 'private sector' to reach goals of innovation.

Along with the integration of 'public data' there is a monumental growth in the data economy of social media and data collecting devices controlled by commercial US digital giants such as Apple, Facebook, Google, Microsoft and Amazon. Here data are shared, yes, but cultural studies scholars have pointed out how the sharing involved is asymmetrical: the citizen 'shares' while the data becomes accumulated behind payment walls (Gehl, 2015), and the use of the data remains proprietary and is essentially kept secret from the data subjects (Crawford et al., 2015). Exclusive access is part of the business model (and therefore public researchers are often barred access). Citizens are made manifest; but the users remain secret. When construed as commercial optimisation, these platforms can enrol all their users as research participants without informing them or others about the research or its results. It is proprietary information. Data might be appropriated from these platforms by so-called 'data scrapers', who find information on social media to generate data profiles for sale (Angwin and Stecklow, 2010), or they can be appropriated by hackers or by national intelligence agencies. The data subjects, however, do not see or control those who pay to use their data.

Digitalised data infrastructures in this way make data available for infinite *potential viewing*: there is always an additional potential reader (Bowker, 2005). Online activity creates data doubles, and the lives of these 'doppelgängers' are inherently promiscuous (Bode and Kristensen, 2015). This form of unwarranted data promiscuity inadvertently puts the individual at risk. As digital infrastructures are set up for data sharing, they always involve multiple points of access. The complexity creates an unspecific form of oblivion and a general inability to grasp who does what with the exchanged data. What can be experienced as a loss of privacy, is more precisely described as a loss of insight into the control of secrecy (Pasquale, 2015). If an old-fashioned letter had been opened, the recipient was likely to find out. However, how will data subjects know who has access to which secrets when digitally stored in complex networks? It is not only the informational aspects of the public–private distinction that are affected by the infrastructural transformations, however, and therefore I now turn to the *institutional* implications.

Institutional implications: 'public' and 'private' moral zones

'Public' and 'private', though clearly not ontologically separate domains, have become entrenched as separate moral zones so that actors viewed as 'private', such as IBM, NETS and KMD, face different political and economic opportunities and responsibilities than do the actors deemed 'public', such as governmental ministries, regions and municipalities. If the public–private distinction as a sense-making device implied incentivising public services (authorities) to operate in the same way as private markets (companies), the performativity of the *moral* zones of the distinction (how people use it to ascribe praise and blame) have proven to operate with reverse effect. Here, 'public' is supposed to be as different as possible from 'private'. It means that profit and blame are distributed unequally, as I will show.

The market value of the 'private' company NETS stems, as described above, from a series of 'publicly' sanctioned monopolies: NemID, E-Boks and digital payment options. Without the Danish authorities making the use of these monopolies easy and sometimes mandatory, there was no 'market', or, it would be considerably smaller. In the case of KMD, its products were developed by and with public authorities, and its customers are primarily these authorities: the income is derived from taxation. Both KMD and NETS, however, have been sold off to equity funds and the profit generated in that process has been appropriated by individuals and companies, not authorities. Like economist Mazzucato argues in *The Entrepreneurial State*, the public sector often makes the investment while the private companies get the fame and the profit (Mazzucato, 2015). When in 2017 an equity fund offered to buy additional shares of NETS, the CEO Bo Nilsson was personally able to cash in more than 620 million DKK (>80 million €), equivalent to the annual turnover of the company. One year later he was once again able to extract more than double that sum. Although the value is clearly derived from its publicly authorised digital monopolies, and though the money to the CEO will have to be paid by the Danish citizens who are obliged by law to use the company services, it is fully legal for this huge amount to become his 'private property' (after just few years of employment). It did give rise to various expressions of envy and critique in the media, but it did not lead to a discussion about making this type of appropriation illegal. It is described, not as theft, but as a bonus.

Above I described how formerly state-owned infrastructures were 'privatised' based on ideas about competence and efficiency associated with for-profit companies. This cultural ideology also seems to overrule and silence the more questionable experiences. For example, when KMD was sold to the equity fund Advent International and when CSC acquired Datacentralen (see above and notes 2 and 3), the sales led to so-called trimming of the 'companies'. Shortly after CSC fired 900 employees in 2013, an unfortunate lack of maintenance of the CPR register led to a leak of 900,000 CPR numbers. The institutional transformation in this way spilled over into informational risks, but it did not lead to media critique of the company's 'competence'. CSC was also supposed to deliver a system to the police (Toft, 2016a) and

another one together with KMD to the tax authorities (Toft, 2016b). Both of them failed. A number of other product failures and a series of notorious data leaks also surround KMD (Toft, 2017). In the case of the system for the tax authorities, some observers estimate the resulting loss of tax revenue for the Danish state to be above 100 billion DKK (13 billion €) (Mølsted, 2017). CSC has recently failed to deliver a functioning update to the national patient register that serves as the backbone of medical research in Denmark, and 2019 is now known as the year of missing data. CSC, however, does not stand to lose.

The *Se and Hør* data leaks from NETS and IBM have been described above too, and it is worth adding that NETS was warned several times about IBM having problems with meeting security standards from 2007 onwards but did not react.[7] Instead, blame is directed at the public authorities when commercial companies fail to deliver according to contract. The National Board of Health Data is held responsible for CSC's failure to deliver a patient registry. The commercial right to profit is uncontested.[8] A company might fail in proving its presumed 'competence'; but when this happens in a public–private partnership it is typically seen as an inability of the public authority to control the process properly.

The point is that when profits can be appropriated by these companies despite their failures, it is indicative of a culture where 'public' and 'private' have become associated with different moral zones. The digital infrastructure is not 'private' or 'public', but the moral zones of economic opportunity and political blame have come to operate in this way. These moral zones shape health data research in various ways by directing blame where little is to be done. It undermines the social sustainability of research that is dependent on data access. Furthermore, actors viewed as 'private', such the pharmaceutical industry, are often ascribed a dubious morality irrespective of what they do. The moral zones of the public–private distinction ascribe doubtful motives to commercial actors in the area of health, irrespective of their actual track history and proven performance. In this way, the moral zones of public and private fundamentally shape the conditions of possibility for health research.

Discussion: searching for new solutions in an age of data promiscuity

This paper has sketched out some of the transformations of infrastructures in Denmark as they have unfolded in the pursuit of a digitalised data economy. First, I have suggested that the transformations have been shaped (though not determined) by the public–private distinction in the sense that the political goals have been to 'privatise' the infrastructure and to model state services on ideas about 'private' markets. Secondly, I have outlined the implications of the infrastructural changes for the informational and the institutional aspects of these infrastructures. In terms of *informational implications*, I have shown how the changes in infrastructure have transformed the conditions under which people negotiate which pieces of information to share with whom. They have not lost an absolute or well-defined form of privacy; but they have lost influence on and insight into the increasingly complex infrastructures and number of stakeholders that now handle their secrets. In terms of *institutional implications*, I have shown how the public–private distinction has been used to justify unequal distribution of responsibilities and profit among agencies depending on whether they are seen as 'public' or 'private'. Monetary flows, work tasks, and informational exchanges between companies and public agencies are fully intertwined, but they operate in different moral zones. The public–private distinction serves not as a simple model *of* reality, but as a model *for* organisational expectations that reinforces this separation of moral zones. In this sense, the distinction has performative effects.

While some observers will see the developments as technology-driven, I have focused on the social urge for data, and on the performative effects of socially engrained notions of 'public' and 'private'. In this way, I have emphasised the social forces at play, not to disregard the agency of the material infrastructure but to direct attention towards the ways in which we think about and organise them (cf. Andrejevic et al., 2015). The point has been to shift the focus from what the distinction means to what it does. As noted with Weintraub (1997) above, we cannot simply erase the distinction, but we can begin to unpack public and private as grand ontological categories, and instead look at who has what at stake in relation to specific aspects (institutional or informational) of the world. When we remember that 'public' and 'private' are sense-making devices with effects on organisation (and not descriptive names for ontological divisions), we can explore more openly the effects of naming and framing something as 'public' or 'private' and move beyond the logics that the categories tend to install. These effects, in turn, are technologically mediated and to trace where data flow involves acknowledging technological agency too.

How might greater awareness of public–private as a fractal distinction then help rethink the problems and concerns outlined above and perhaps pave the way for new solutions? First, it can make us see how the GDPR draws upon the public–private distinction in unfortunate ways. The GDPR gave individuals 'rights' to access, delete and sell (through the 'portability principle') data on themselves, but the very notion of 'portability' here reveals how this approach to privacy as matters of private property is as much about facilitating as retraining data exchange. It reifies ideas about data as 'private', discrete entities available for commercial exchange, as foreseen already by Lyotard (1984) in *The Postmodern Condition* (see also Prainsack, 2019).[9] However, data are multiple and unstable objects and never just data on one person or one thing. Health data might be data on a patient's disease, but it is also data on the treating physician, the hospital, and the laboratory that delivered a test. Therefore the use of the property model, 'it is the patient's data', does not work. When information is datafied, it is to multiply its uses. Even when data are treated as property by platforms conjuring data, these platforms monetise on exclusive control with data based on the ability to reinvent the meaning and purpose of data uses (van Dijck et al., 2018). Data are relational and can become data on whatever a researcher makes the subject of investigation in the course of exploring correlations. Digital integration means that data can move more easily across domains. Health data can be used by the social services, for credit assessment, or potential employers and vice versa. In consequence, it is the basic premise of data to be promiscuous.

Individual control is an inadequate measure for ensuring privacy interests. Just as an ethics of monogamy fared better as a technology of blame than as a means of protection in relation to sexual promiscuity, responsibilisation of the individual does more to protect the institutions and companies thriving on data than the individuals who are asked to hand over data as a precondition for entering social relations and accessing services. As shown above, individual access to health data via public platforms can cause loss of control. Similarly, the individual 'right' to enter agreements with commercial digital platforms can imply surrendering rights rather than gaining them. As a consequence of the conflation of the informational and institutional meaning of 'private', people are asked to click to accept terms of agreement or cookies, or sign informed consent sheets. It is not even practically doable to read all of the 'terms of agreements' that citizens need to accept to live a digital life (McDonald and Cranor, 2008). It really is time to look for new legal tools.

Instead of proclaiming an individual right to privacy and suggesting it should be protected with property rights that cannot be upheld anyway, it is time for a much more thorough political debate about who may use which data for what purposes. Simmel encouraged scholars to elucidate the politics of the secret and the manifest, and digitally mediated research infrastructures are good places to do so. If we relinquish the conflation of privacy and property, it might become easier to imagine new and more collective approaches (Mittelstadt and Floridi, 2015; Taylor, 2017). There might, for example, be room for inventing new types of data custodians appointed by democratic bodies with well-defined responsibilities and entitlements to negotiate data uses on behalf of citizens—someone with a much more defined political role than what the GDPR calls a data protection officer. Moreover, how about focusing not on the right of the individual to exert privacy, but on shared legal safeguards that can step in *when* leakages happen? Prainsack (2017) suggests harm mitigation funds. There is also room for developing new bans on unproductive or harmful data use. There are elements of the market for data profiling that could become subject to plain prohibition (e.g., the so-called 'consumer reporting companies' that profit from selling data on people to recruitment agencies, insurance companies, and companies wanting to do price discrimination). The point is to transgress the public–private logic and its implicit divisions and conflations in order to consider collective solutions that might better mitigate the potential harms produced by data intensification.

To begin inventing new logics that move beyond the property model, it is necessary to realise that there never were such a thing as 100% 'private' data. It is an old insight. Simmel noted that 'writing is opposed to all secrecy (…) it involves an unlimited, even if only potential, "publicity"' (Simmel, 1950b: p. 352). Geoffrey Bowker describes the database as indicative of an 'epoch of potential memory' where 'the question is not what the state 'knows' about a particular individual, say, but what it can know *should the need ever arise*' (Bowker, 2005: p. 30, emphasis original). With digitalisation, this potentiality is augmented to an unprecedented degree. It stretches far beyond 'the state'. Digitalised data infrastructures are devices for *potential viewing*. People use the expression 'my data', but 'my data' is an oxymoron; data are always potentially shared. Otherwise they would not be data. It is an ontological premise for digital data to be 'potentially viewed' by unexpected users: data infrastructures are constructed with some sense of promiscuous data dating in mind. Citizens leaving data traces as a precondition for pursuing a normal life should therefore not be held accountable for what they have few options for controlling.

Finally, there is a need to rethink how the public–private distinction legitimises the distribution of revenue. The revenue generated from state-sanctioned tools such as E-Boks could remain in the hands of either the authorities or in the pockets of the citizens, for example through bans on extravagant bonus programmes. There is no going back in time, however, and no point in longing for a time when infrastructures were run by the authorities. Rather, the point is for social science critique to stop projecting clarity back in time as if our current moment has simply lost a division between public and private that used to work. Critical data studies must work with data science to create more socially robust infrastructures (Neff et al., 2017). Our current problems do not reflect a sudden 'blurring' of previously distinct ontological domains. They reflect an inability to address the inequalities and lack of justice right in front of us.

If scientific research is to retain its legitimacy in the wider public, it should not abstain from using data. Rather, it must use data to generate insights into such inequalities and other concrete problems. Research can help find ways of mitigating them. The social sustainability of health research will not be ensured with intensified informed consent demands, but with research serving collective interests without putting individuals at risk.

Data availability

For research ethical reasons, it is not possible to share the interview transcripts in full.

Received: 24 March 2020; Accepted: 30 June 2020;
Published online: 16 July 2020

Notes

1 Related to these two aspects, but cutting across them, some also use the distinction to talk about *spheres* (a public sphere for open debate and a private sphere of conversation among family and friends) or *spaces* (public city spaces subject to formal rules versus private secluded spaces for unregulated interaction) (Sennet, 2017; Sheller and Urry, 2003).

2 While there are many similarities between the Nordic countries, there are also significant differences in way the public–private distinction has played out. Sweden has generally opted for legitimacy through transparency, whereby for example tax information is presented as public, while Denmark and to some extent Norway presents the authorities as guardians of secrets.

3 The US technology giant CSC bought the public Danish IT supplier *Datacentralen af 1959* and thus acquired access to the governmental sector. Since 2017, the Danish branch has taken the name DXC Denmark, but for the sake of simplicity I use CSC throughout here.

4 KMD, or Kommunedata as it was originally called (*Eng.* Municipality Data), was established in 1972 and originally a publicly owned service provider. It was privatised in 2009 but it still holds the contract for processing most public accounting and salary systems.

5 See also a list of data leakages involving the CPR number: http://www.dr.dk/nyheder/viden/tech/her-er-ti-moegsager-det-danske-cpr-nummer, last accessed Nov 15, 2017.

6 The national credit card system in many ways protects credit card information better than in, for example, the US (O'Neil, 2016), so my point here is merely that the commercial infrastructural guardians, NETS and IBM, are not facing any serious reproach for the leakages.

7 https://www.computerworld.dk/art/230922/igen-og-igen-og-igen-saa-mange-advarsler-har-nets-faaet-siden-2007, last accessed Nov 15, 2017.

8 The same mechanism unfolded in 2017 in Sweden as it turned out that IBM had outsourced the maintenance of confidential Swedish registers to cheap Czech suppliers without providing security checks. This ended up with two government ministers losing their jobs for failing to detect the breach.

9 The GDPR should also be acknowledged for having raised awareness of privacy interests and for trying to hold institutions accountable for data security and technical safeguards. Technical safeguards are very important indeed; they are just not enough.

References

Andrejevic M (2013) Infoglut. How too much information is changing the way we think and know. Routledge, London
Andrejevic M, Hearn A, Kennedy H (2015) Cultural Studies of data mining: Introduction. Eur J Cult Stud 18(4–5):379–394
Angwin J, Stecklow S (2010) 'Scrapers' dig deep for data on web. Wall Street J. December 10
Ariès P (1989) Introduction: passion of the renaissance. In: Chartier R (ed) A history of private life, vol 3. The Belknap Press of Harvard University, Cambridge, pp. 1–11
Bode M, Kristensen DB (2015) The digital doppelgänger within. A study on self-tracking and the quantified self movement. In: Bajde D, Canniford R (eds) Assembling consumption. Routledge, Oxford
Bowker GC (2005) Memory practices in the sciences. The MIT Press, Cambridge
Crawford K, Lingel J, Karppi T (2015) Our metrics, ourselves: a hundred years of self-tracking from the weight scale to the wrist wearable device. Eur J Cult Stud 18(4–5):479–496
Danske Regioner (2017) Sundhed for alle. Vision for et bæredygtigt sundhedsvæsen. Danske Regioner, Copenhagen
Finansministeriet (1983) Redegørelse til Folketinget om Regeringens program for modernisering af den offentlige sektor. DenOffentlige.dk, Copenhagen
Forskningsministeriet (1994) Info-Samfundet år 2000. Forskningsministeriet, Copenhagen
Frank L (2000) When an entire country is a cohort. Science 287(5462):2398–2399
Frank L (2003) The epidemiologist's dream: Denmark. Science 301(5630):163

Frost and Sullivan (2017) Digitilization in healthcare: emergence of digital health portals. Frost and Sullivan, Mountain View
Gal S (2002) A semiotics of the public/private distinction. J Femin Cult Stud 13(1):77–95
Geertz C (1973) The interpretation of cultures. Basic Books, New York, NY
Gehl RW (2015) Sharing, knowledge management and big data: a partial genealogy of the data scientist. Eur J Cult Stud 18(4–5):413–428
Hoeyer K (2016) Denmark at a crossroad? Intensified data sourcing in a research radical country. In: Mittelstadt BD, Floridi L (eds) The ethics of biomedical big data. Springer, Dordrecht, pp. 73–93
Hoeyer K (2019) Data as promise: Regonfiguring Danish public health through personalized medicine. Soc Stud Sci 49(4):531–555
Johansen VF, Andrews TM (2016) On challenges to the private-public dichotomy. Soc Theor Health 15(1):66–83
Leonelli S (2016) Data-centric biology. A philosphical study. The University of Chicago Press, Chicago
Lindstrøm M, Pedersen NK, Hougaard IB, Lynnerup MCRG (2017) Sundhed i skyen: Et kig ind i den digitale fremtid på sundhedsområdet. Mandag Morgen, Copenhagen
Lyotard J-F (1984) The postmodern condition. A report on knowledge. University of Minnestoa Press, Minneapolis
Mazzucato M (2015) The entrepreneurial state. Debunking public vs. private sector myths. Anthem Press, Cambridge
McDonald AM, Cranor LF (2008) The cost of reading privacy policies. J Law Policy Inform Soc 4:540–565
Mittelstadt BD, Floridi L (2015) The ethics of big data: current and foreseeable issues in biomedical contexts. Sci Eng Ethics 22(2):303–341
Mølsted H (2017) Skat godt på vej med EFI-afløser: Det skal gå stærkt, for milliardgælden vokser. Version 2, Apr 4
Neff G, Tanweer A, Fiore-Gartland B, Osburn L (2017) Critique and contribute: a practice-based framework for improving critical data studies and data science. Big Data 5:85–97
Pasquale F (2015) The black box society-the secret algorithms that control money and information. Harvard University Press, Boston
O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy. Allen Lane, London
Organisation for Economic Co-Operation and Development (OECD) (2013) OECD reviews of health care quality: Denmark. Executive summary. OECD, Pari
Prainsack B (2017) Personalized medicine-empowered patients in the 21st century? New York University, New York, NY
Prainsack B, Krutzinna J, Floridi L (2019) Data donation: how to resist the iLeviathan? The ethics of medical data donation. Springer, Dordrecht
Retten i Glostrup (2016) Dom i Se og Hør-sagen. Courthouse Glostrup, Glostrup
Ruckenstein M, Schüll ND (2017) The Datafication of Health. Ann Rev Anthropol 46:261–278
Ruppert E, Isin E, Bigo D (2017) Data Politics. Big Data Soc 4(2):1–7
Sætnan AR, Lommel HM, Hammer S (2011) Introduction. By the very act of counting–The mutual construction of statistics and society. In: Sætnan AR, Lomell HM, Hammer S (eds) The mutual construction statistics and society. Routledge Talor and Francis Group, New York, NY, pp. 1–21
Sennet R (2017) The fall of public man. W.W. Norton and Company Inc, New York, NY
Sheller M, Urry J (2003) Mobile transformation of 'public' and 'private' Life. Theor Culture Soc 20(3):107–125
Simmel G (1950a) Secrecy. In: Wolff KH (ed) The sociology of Georg Simmel. The Free Press, Glencoe, pp. 330–345
Simmel G (1950b) The secret society. In: Wolf KH (ed) The sociology of Georg Simmel. The Free Press, Glencoe, pp. 345–376
Tarkkala H, Helen I, Snell K (2019) From health to wealth: the future of personalized medicine in the making. Futures 109:142–152
Taylor L (2017) Safety in numbers? group privacy and big data analytics in the developing world. In: Taylor L, Floridi L, van der Sloot B (eds) Group Privacy. Spring, Cham, pp. 13–36
Toft D (2016a) Politiet røvrendes med dybt forældet it-og millionerne fosser ud af kassen. DR.dk, October 25
Toft E (2016b) Gælden til det offentlige vokser med milliarder: LA kalder det skattelotteri. DR.dk, April 9
Toft E (2017) It-hul i pladsanvisningen har givet adgang til cpr-numre og navne i 12 år. DR.dk, June 15.
van Dijck J (2014) Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. Surveill Soc 12(2):197–208
van Dijck J, Poell T, De Wall M (2018) The platform society. Public values in a connective world. Oxford University Press, New York City, NY
Wadmann S, Hoeyer K (2018) Dangers of the digital fit: Rethinking seamlessness and social sustainability in data-intensive healthcare. Big Data Soc 1–13.
Warren SD, Brandeis LD (1890) The right to privacy. Harvard Law Rev 4(5): 193–22
Weintraub J (1997) The theory and politics of the public/private distinction. In: Weintraub J, Kumar K (eds) Public and private in thought and practice. The University of Chicago Press, Chicago, pp. 1–42
Winthereik BR, van der Ploeg I, Berg M (2007) The electronic patient record as a meaningful audit tool. Accountability and autonomy in general practitioner work. Sci Technol Hum Values 32(1):6–25

Acknowledgements

I would like to thank the interviewees who have been willing to share their time and experiences with me, and the reviewers for improving the text. I would also like to thank Tamar Sharon and Minna Ruckenstein as well as members of our research group on data practices for comments on earlier versions. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 682110).

Competing interests

The author declares no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.H.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.