Body data—data body: Tracing ambiguous trajectories of data bodies between empowerment and social control in the context of health

Astrid Mager* und Katja Mayer**

Abstract

A plethora of health apps and tracking devices is used around the globe to measure, store, and process body data. In this article, we use various approaches from the fields of science and technology studies (STS), surveillance studies and medical sociology to grasp and theorize these global trends of body datafication in health-related contexts. We (re)introduce the post-digital concept of the data body as the intersection of online and offline, individual and collective, private and public aspects, emphasizing the entanglements of the physical body from its data-dimensions and its situatedness between empowerment and social control. We conclude by discussing aspects of ownership, care, and control of digital data bodies and how both individuals and society may cope with them in the future.

Keywords: quantified self, digital health, data governance, surveillance medicine, patient empowerment

Körperdaten – Datenkörper. Auf den Spuren mehrdeutiger Reisen von Datenkörpern zwischen Empowerment und sozialer Kontrolle im Gesundheitsbereich

Zusammenfassung


Schlagwörter: Quantified Self, digitale Gesundheit, Datenpolitik, Überwachungsmedizin, PatientInnen-Selbstermächtigung
1. Introduction

“The data body has two primary functions. The first purpose serves the repressive apparatus; the second serves the marketing apparatus. The desire of authoritarian power to make the lives of its subordinates perfectly transparent achieves satisfaction through the data body. Everyone is under permanent surveillance by virtue of their necessary interaction with the marketplace. Just how detailed data body information actually may be is a matter of speculation, but we can be certain that it is more detailed than we would like it to be, or care to think.” (Critical Art Ensemble 1997: 145)

We are currently witnessing a ubiquitous privatization of personal health-related data. Myriads of health apps and tracking devices are used around the globe to measure, store, and process body data. Fitness apps trace running routes and other physical activities, medical devices measure blood sugar levels, sleeping habits, or mood swings, and nutrition apps track daily intakes of food or suggest how to lose weight. Additionally, various companies offer online genetic testing on the basis of saliva samples. Even more radically, biohacking communities have started to develop and make available tools and bioinformatics resources for biological engineering. These are just some examples that show how digital tools and technology are used to monitor and enhance the body. They illustrate how hardware, software, and bodily functions merge into a machinery neatly surveilling, storing, and optimizing the physical body and its context. In this process, body data is being transformed into data bodies. Building on Lupton’s (2018) concept of “human-data assemblages” and Lyon’s (2005) notion of “data double,” we define data bodies not as mere digital counterparts of physical bodies, but rather as an indispensable socio-material coupling of data and bodies. The data body should hence be seen as co-configured from both online and offline data, behavior, and body signals. In this article, we demonstrate how this double notion of body data/body data as both socio-material entity and set of practices allows us to overcome traditional dichotomies. Following the sociopolitical objective underlying this article, we further make the communal aspects of digital health data visible and open them up for policy actions, especially in this age of increasing privatization of health-related data and fundamental transformations of governments and public health sectors.

The following questions will guide our analysis:

• What is driving the increasing quantification of body and health?
• How are body data transformed into data bodies?
• How do they oscillate between empowerment and social control?
• How can we deal with increasingly detailed data bodies in the future?

To answer these questions, we draw on literature from the fields of science and technology studies (STS), surveillance studies and medical sociology. Moreover, by bringing these schools of thought together, we aim to enrich the predominantly functionalistic discourse on smart technologies or smart services (Lim & Maglio 2018). In literature discussions of services that involve intensive data and information interactions in the health domain—e.g., management of patient health, gathering of personal health records, devices for health care and monitoring—the main focus is on value creation and how to improve human health (Maglio, Kwan, & Spohrer 2015). Contrary to this body of work, we focus on data bodies and their sociopolitical dimensions. Defining data bodies as the socio-material coupling of data and bodies enables us to broaden the perspective and challenge the engineering-centric logic underlying much of the smart services literature that currently informs policy debates (Maglio 2015). Moreover, we aim to further complicate the picture by showing how data bodies constantly oscillate between, blur, and break up traditional differentiations between growing capacities of action and new forms of discrimination, between autonomy and heteronomy, but also between the individual and the collective. In this analysis, we use the term “data body” because it signifies the inseparability of the physical body from its “virtual,” “semiotic,” “sign” dimensions (Krämer 2008), while also opening it up and showing its multi-faceted dimensions going beyond merely dystopian characteristics, as suggested in the quotation above. In a post-digital understanding of performativity, realities are shaped online and offline, bodies can be regarded as intersections of the individual and the collective, the private and the public, the flesh and the machine (Apprich 2018; Berry and Dieter 2015; Critical Art Ensemble 1997).

In the first section, we trace the birth of the data body in twentieth-century “surveillance medicine” (Armstrong 1995) and discuss various driving forces behind the increasing quantification of health and the body since then. In the second section, we conceptualize the data body as the socio-material coupling of body and data by drawing on actor-network theory.
(Latour 2005; Law and Hassard 1999) and new materialism (Barad 2003; Haraway 2003, 2008), socio-material perspectives in the medical field most particularly (Lupton 2016, 2018). We then discuss ambiguous trajectories of, and intrinsic tensions within, data bodies by focusing on quantified self-communities and practices of socioeconomic optimization and social scoring. In this analysis, we show, juxtapose, and compare different motivations, benefits, and dangers related to (self-)surveillance. We follow the creation, archiving, processing, sharing, categorizing, and mobility of data bodies on their travels between mundane practices of self-surveillance, corporate desires, and institutional interests. To conclude, we discuss implications of this analysis in regard to ownership, care, and control of digital data bodies and how individuals and society may cope with them in the future.

2. The birth of the data body

Surveillance is not a new phenomenon in the medical field—quite the contrary. Armstrong (1995) traces “surveillance medicine” back to the early twentieth century. Building on work identifying different medical cosmologies including bedside medicine, hospital medicine, and laboratory medicine (Ackernacht 1967; Jewson 1976), Armstrong coined the notion of surveillance medicine to capture the shift from individual treatment to statistical measures and public health initiatives. With its roots in the monitoring of children during the twentieth century, the author describes height and weight growth charts as paradigmatic images of surveillance medicine. Height and weight charts draw curving lines representing growth trajectories. The individual trajectory, however, can only be measured in a context of general population trajectories. The same applies to socio-medical surveys introduced during World War Two and extensive screening programs introduced after the war to monitor individuals’ health. This “machinery of observation” (Armstrong 1995) has profound implications for the very nature of medicine and illness. The physical body is no longer the mere locus of illness, but “illness begins to leave the three-dimensional confine of the volume of the human body to inhabit a novel extracorporal space” (Armstrong 1995: 395). “Medical surveillance hence would have to leave the hospital and penetrate into the wider population,” as Armstrong (1995: 398) concludes. Surveillance medicine may thus be seen as giving birth to the data body that is supposed to be monitored and normalized. What constitutes “the normal,” however, becomes problematized in the course of the advent of new measuring instruments. Moreover, identity is reconfigured due to new techniques of observation that blur the boundary between health and illness. A new category is introduced: the healthy body at risk. “Identity then begins to crystallise in a novel temporal and multidimensional space whose main axes are the population—within which risk is located and from which risk is calculated—and a temporal space of possibility” (Armstrong 1995: 403).

Prainsack (2017) speaks of blurring boundaries between health and lifestyle in this context. With the rise of health apps, smart watches, and so forth, practices of data collection for medical purposes overlap with data collection for wider purposes, leading to “data hyper-collection” (Prainsack 2017: 49).

Since risk calculation is an essential part of the insurance sector, it comes as no surprise that insurance businesses also emerged as central drivers behind the increasing quantification of the human body and life. The insurance sector has been at the forefront of the development of statistics both as a scientific discipline (Desrosieres 2002) and as a social technology (Mayer 2012). Bouk (2015)—among others—has studied the historical entanglement of life insurance companies and the development of medical testing and routine examination in the USA and demonstrates their power in co-shaping how people think about and treat their bodies. Moreover, new information providers, platforms, gatekeepers, and, ultimately, new markets have emerged. Access to health information, monitoring devices, community platforms, online genetic testing, and so forth are mainly provided by big corporations, often based in Silicon Valley. The desire to monitor and measure data bodies hence no longer stems from medical and public health institutions, as in classical surveillance medicine, but rather is now driven by commercial actors. Saukko (2018) describes “digital health” as a new medical cosmology driven by commercial digital health platforms and devices. Commercial health apps, platforms, and communities build on features of surveillance medicine, such as providing consumers with information on what is normal and how to change their behavior, which indicates that they are more than just smart service technologies. They are fundamentally changing how people interact with health expertise and make sense of their bodies. Embedded in digital culture and marketing, these platforms and apps make digital health seem less paternalistic and more open-ended than health advice propagated by expert organizations.
or hospitals. While public health institutions have provided the ground for self-surveillance, big corporations are riding on the wave of surveillance medicine with private means.

Finally, individuals themselves have contributed to the increasing quantification of body functions and human lives. Ordinary people became fascinated by the shift from "a subjective to a numeric approach to evaluating weight" (Czerniawski 2007: 273). Igo (2007) links this fascination with the emergence of a "mass society" and the new establishment of self-awareness along statistical lines. It was already possible to accurately predict future events (e.g., mortality rates) with the invention of "probability" that was then backed by the "normalization" of human beings. Hacking (1990) describes this historical shift from deterministic causation to the idea that phenomena are a distribution of characteristics as a new style of reasoning, one that is still at work today and firmly embedded in data practices everywhere. In a post-demographic understanding of data, as Mackenzie (2016: 116) illustrates,

"individuals appear not simply as members of a population (although they certainly do that), but themselves as a kind of joint probability distribution at the conjunction of many different numbering practices. If individuals were once collected, grouped, ranked, and trained in populations characterized by disparate attributes (life expectancies, socio-economic variables, educational development, and so on), today we might say that they are distributed across populations of different kinds that intersect through them."

Biological and social normality is coupled and constructed with a range of characteristics over populations. Thus, the individual represents an intersection of diverse populations and is defined in regard to statistical standards of measuring and analysis that always imply a form of totality. In that sense individuals are not singular, they are regular.

3. Data bodies as socio-material coupling of body and data

After having discussed various drivers of quantification in the health-related context, we now turn to defining data bodies as socio-material coupling of body and data. In doing so, we situate data bodies in approaches from classical actor-network theory (Latour 2005; Law and Hassard 1999) and new materialism (Barad 2003; Haraway 2003, 2008); socio-material perspectives from the medical context more specifically (Lupton 2016, 2018). Actor-network theory (ANT) is a material–semiotic approach that enables us to analyze social reality as being shaped in a network of "materially and discursively heterogeneous relations that produce and reshuffle all kinds of actors including objects, subjects, human beings, machines" (Law 2007). It hence serves as a valuable tool for symmetrically analyzing human and non-human entities and their heterogeneous entanglements (Latour 2005; Law and Hassard 1999). Moreover, it makes it possible to understand the dynamic character of each of the elements involved in the actor-network that shapes reality, since it defines identity not as given but rather as enacted and stabilized by the relational effects of the network (Law 2007). Hence actors do not have a predefined identity; rather their identity is made and remade dynamically by the actor-network around them. Through the analytical lens of ANT, we define data bodies as being co-shaped by human and non-human entities including human behavior and body signals, as well as software, hardware, and large-scale, often corporate infrastructures.

Within the overall framework of ANT, we draw on contemporary research in new materialism—socio-material perspectives more specifically—to define data bodies as the socio-material coupling of data and bodies rather than digital counterparts of physical bodies. In this perspective actors are intrinsically interwoven with the environments in and through which they move (Barad 2003; Haraway 2003, 2008). Drawing on socio-material perspectives, Lupton (2018) coins the notion "human-data assemblages" to discuss tight entanglements of humans and non-humans in regard to quantification in the medical and health-related field. Digital data assemblages enable us to bring together heterogeneous elements including humans, devices, software, data, space, and time and perceive them as co-shaping each other: Just as humans cannot be separated from their environments, the body cannot be separated from its data. "From the sociomaterial perspective, data about humans are always part of each other and emerge together. Just as it can be claimed that authors and books write each other (Barad, 2007: x [sic]); it can also be asserted that people and their data make each other" (Lupton 2018: 5).

To capture the hybrid, unstable, and generative character of body data, Lupton (2016, 2018) introduces the term "lively data." Using Haraway's notion of "compansion species" (2003), Lupton (2016, 2018) speaks of digital data as "lively data" since they may be seen as lively combinations of nature and culture. They are lively in terms of containing information about human life, but
also in terms of having a life of their own. Resonating with central ideas in ANT, Lupton (2018: 6) further argues that data have “thing-power” (Bennett 2004) in the sense that they can shape people’s embodied responses and actions, their sense of selfhood, and their relationships with other people and things. Moreover, personal digital data can take on different meanings in different contexts by being cleaned, combined, and recombined with other data sets. Lupton (2018) hence concludes that personal digital data possess biovalue. They may be seen as a “new type of human remains, one that is potentially open to a multitude of repurposing and reconfiguring, leading to many kinds of value for a diverse range of actors” (Lupton 2018: 6). In these processes of combination and recombination, body data are transformed into data bodies leading a life of their own. Which lives they lead on their complex trajectories between individual empowerment and social control will be elaborated in the next section. We will further discuss how value is created, who benefits in what contexts, and what politics and “biopower” (Foucault 1998: 40) data bodies entail. Lyon (2005: 27) argues “[…] the data doubles, created as they are from coded categories, are not innocent or innocuous virtual fictions. They affect eligibilities for credit or state benefits and they bestow credentials or generate suspicion. They make a real difference. They have ethics, politics.” And in that sense, they are no longer “doubles;” they are not a kind of mask through which bodies speak, but they are intrinsically interwoven with bodily practices and biopolitics. With the term “biopower” Foucault (1998: 40) refers to “an explosion of numerous and diverse techniques for achieving the subjugations of bodies and the control of populations.” Consequently, we are interested in this socio-material coupling of body and data in all its different forms, ranging from wearable technology via analytic practices to changing human behavior and large-scale surveillance of populations.

1 “And, while Foucault is somewhat imprecise in his use of the terms, within the field of biopower, we can use the term ‘biopolitics’ to embrace all the specific strategies and contestations over problematizations of collective human vitality, morbidity and mortality; over the forms of knowledge, regimes of authority and practices of intervention that are desirable, legitimate and efficacious” write Rabinow and Rose (2006:197) in response to the use of the term by Foucault.

2 Even though Foucault did not use the term “population” in a statistical sense in this quotation, we can stretch the meaning of his concept further.

4. Data bodies between empowerment and social control

4.1 Self-tracking as caring and collective experience

Based on our interest in the different types of coupling of data and body that we witness today, it is certainly worth looking closer at the cluster of phenomena that has been labeled “The Quantified Self,” which refers to tracking the physical, environmental, biological, and behavioral aspects of everyday life. As Lupton (2016a: 104) notes, “traditional self-tracking practices have included age-old strategies such as journaling and diary-keeping. However, the recent focus on monitoring the self in both popular forums and the academic literature centers on using digital technologies.” These digital technologies co-shape the very definition of the “self” and reconfigure both notions of embodiment and social relationships. In 2007 as Gary Wolf, one of the founding fathers of the movement, recalled, new tools like “life logging, personal genomics, location tracking, biometrics […] were being developed for many different reasons, but all of them had something in common: they added a computational dimension to ordinary existence” (Wolf 2011). As illustrated in this quotation, quantification is used here to describe a broad range of types of self-tracking, referring to various wearable and sensing technologies. Typical wearables, such as smart watches, can measure and analyze the heart rate, blood pressure, steps taken, calories burned, time spent sleeping or exercising, and so on based on a combination of automated data collection and self-reporting. Since 2007 the movement, which operates with the slogan “Self-Knowledge through Numbers,” has grown and the topic of self-tracking has gained momentum in a broad range of societal sectors and attracted attention from both public and private actors. It is reported that the global community comprises more than one hundred local groups, the largest ones with more than 1,500 active members.3 The 2018 Quantified Self conference3 featured over 80 presentations, workshops, and tool demonstrations, and showcased the expertise of

3 In the quantified-self context we often see quotations or references to Foucault’s concept of “care for the self” (2012).


many local meet-up groups from Western or emerging countries. At the dawn of the big data hype, enthusiastic voices from the health-care sector imagined an “unprecedented super-convergence” (Topol 2013) of knowledge based on ubiquitous connectivity and data gathering to the highest granularity, which would bring an age of democratization and improvement of health care due to dramatic power shifts from caregivers to patients. Even though we did not witness these disruptive changes, industry is reporting a continuing increase in the sale figures of wearable devices, such as health and fitness trackers. However, self-tracking practices should not be seen as one-dimensional and monolithic, but rather as highly diverse and context-specific. Consequently, the gains and drawbacks of self-tracking practices vary according to the respective context, as discussed below.

For many patients with diabetes or similar conditions, studies have shown improvements based on the day-to-day frequency of self-monitoring, e.g., by wearing sensors and conducting blood glucose tests (Miller et al. 2013). In such contexts the use of medical self-tracking technologies is already widespread and has a long tradition. Patients manage their conditions and perform “patient work” (Strauss et al 1982; Mathieu-Fritz and Guillot 2017) in accordance with their doctors and other caregivers. In such settings the patients have long been “actors in the professional construction of illness and the medical division of labour” (Baszanger 1986: 6). Moreover, patients not only have the capacity to carry out self-care, e.g., by administering the required dose and type of insulin, but their expertise is also increasingly acknowledged in the patient–doctor relationship. Furthermore, in such settings the body is not mainly treated as a passive object of measurement, but the patient’s agency is enhanced by the coupling of body and data. In the context of chronic illness, studies also show that new mobile technologies such as health apps are not just adopted or easily appropriated. On the contrary, users become interested in the data flows and develop unforeseen practices, such as collecting their data but then refusing to share these data if they cannot take part in the definition of the objectives underlying a data analysis system (Piras and Miele 2017; Rabeharisoa et al. 2013).

When analyzing the many different types of interactions with health-care monitoring equipment for self-tracking, we become aware of how their users are developing their own strategies to incorporate them into the treatment process. Moreover, we can see that rates of adoption are closely linked to the trust in the applications and the biopolitical practicalities of how best to manage daily life with a medical condition. Indeed, such tracking technologies can offer new degrees of freedom to patients who need personalized care and have struggled with the health system because it has long been biased toward data from overrepresented populations, such as white Western adult male bodies, for example (Magnet 2011; Ajana 2013). Even though self-tracking can be viewed critically in light of its contributions to standardization, indeed normalization, we should not forget that it is also breaking up rigid classification systems by introducing a broader perspective and more granular classifiers, even more inclusive regularities. However, it remains necessary to closely observe how current monitoring and classification systems will evolve and to whose benefit.

Lupton (2016a) differentiates five overlapping modes of self-tracking, which all to a certain degree can be found in the settings described in this article: private (personal and non-shared), pushed (incentivized by another actor, e.g., patient self-care programs or insurance plans), imposed (e.g., RFID tracking), exploited (as seen in the recent Facebook scandal), and finally communal modes of tracking (with an emphasis on data sharing, e.g., health data platforms). It is this communal dimension of “digital care” and self-tracking that we would like to address henceforth: digital biosociality. Biosociality—a concept introduced by Rabinow (1996)—marks new types of sociality that emerge when knowledge about diseases and human bodies changes. In contrast to historical accounts of sociobiology, Rabinow (1996) documents the formation and empowerment of collectives and identities, as well as the options for new types of expertise.

There is more to self-monitoring—or “self-metrating”—than just treating it simply as an individualized service or a “nurselike application of technology” (Singer 2015). Whereas most scholarly literature and commentators focus on the individual scope of tracking, only a few studies highlight the communal effect and potentials that are associated with trackers in connected realms. “Personal data are ideally suited to a social life of sharing. You might not always have something to say, but you always have a number to report,” said the co-founder of the Quantified Self community (Wolf 2010). Gamification and platforms for sharing fitness data have attached a competitive value (another powerful nudging instrument) to the tracking technologies. They have co-shaped communities of
tracking that share not only their concerns or encouragement but also the facts about their bodies and their surroundings. Sharing personal tracking data is part of a general trend toward new forms of biosociality that widen the scope of socialization via new mobile online interfaces that link biotechnologies, humans, organizations and co-create new communities (Hagen 2010). Platforms like “Patients Like Me” want to counter a culture of distrust (toward pharmaceutical research) with “data philanthropy” (UNGP 2009; Pawelke and Tatevossian 2013). With data philanthropy the United Nations Global Pulse initiative encourages strategic partnerships between private and public entities for the advancement of biomedical research and a more holistic understanding of public health. Based on understanding the diverging potentials of “digital biosociality,” Jordan and Pfarr (2014) call for aspiring to the “Quantified Us.” They differentiate between three types of immediate benefit of self-tracking by numbers: 1) gratuitous behavioral data, 2) motivation for self-improvement, 3) enhanced decision-making. However, all these aspects only become meaningful when seen through the lens of a community and through constant exchange.

Whereas the “Us” remains to be inscribed as locus between small and big data (Ajana 2017) where sense-making and collaboration happens, this analytic space is already in the crosshairs of insurance companies and health-related business models. Besides setting the prices of premiums and analyzing their risks, insurance companies and fitness platforms alike encourage fitness team competitions in companies. However, there is more to the communal tracking of “social fitness” (Lupton 2013b) as it also nurtures the desire to be part of a community. Sharing data—or the “donation of data” (Sharon 2017)—is regarded as an act of solidarity and data become a medium for connecting with others—quite the contrary to the stereotypical pigeonholing of self-trackers as narcissistic egomaniacs. This does not always require online platforms, as the regular meetups of the Quantified Self communities demonstrate. Over the years a robust structure has evolved, which also features very critical discussions of self-tracking practices, e.g., in “show and tell” presentations, which also often address the messiness of the socio-technical settings or privacy concerns. Even though those communities may seem “particularistic” and represent only “narrow” forms of solidarity, as Sharon (2017) has shown, they could establish powerful voices and infrastructures to draw attention to issues in health care in general. Sage Bionetworks’ is one example of a patient community platform that allows the management and curation of personal health data for collaborative data analysis.

4.2 Objectification, privatization, and commodification

While self-tracking tools provide new forms of individual freedom and sociality, they may also be regarded as perfect nudging instruments and triggers (similar to the weight scale) for individuals to take responsibility for their own health (Ajana 2017). This responsibility for health is often put in the context of similar strategies of self-governance. As Lupton (2015) argues, these data-driven practices of health management comply with a broader trend of neoliberal politics that shifts issues of personal governance away from institutions and to the individual. It coincides with a “projectification” of human life that is dedicated to constant self-optimization, self-development, and investment (Ajana 2017). In such a context, the “self” figures as a body of knowledge that can be discovered, shaped, and governed. This notion of functionalistic self-quantification has been widely criticized as a neoliberal ideal by authors such as Lupton (2016b) or Moore and Robinson (2016). Wolf (2010) explains that one particular foundation of the Quantified Self movement is the desire to overcome human fallibility: “We make errors of fact and errors of judgment” (Wolf 2010). Ajana (2017) discusses the implications of framing the relation of the body and the self as a direct connection between technology and truth: “Data emerging out of bodily quantification are believed to reveal some kind of ‘objective truth’ about the self-tracker” (Ajana 2017: 4). This liberation from error and intuition, this objectification is regarded as “free choice” that even comes with the promise of reward, e.g., a bonus in the user’s health insurance plan. The tracking self is now in control; furthermore, the self can be objectified and regulated according to whatever norms are either built into the measurement system or are worth aspiring to, for example going to bodily extremes (in terms of body shape or fitness). This self can also be commodified into the vast market

---


options for better self-awareness and optimization of one's well-being.

With digital profiling making the gathering of large amounts of personal data and thus multidimensional health data a reality and potentially speeding up the processes “from bench to bed,” sociometric industries (Mayer 2009) like Google, Tencent, and Apple are becoming more and more involved in the collection and analysis of health data. They are developing into data monopolies, not only by controlling the methods and instruments but also by owning the data and the profiles and being the gatekeepers of access to personal data flows. Besides their business models that exploit digital health practices and personal data for personalized advertising, they also hold more sensitive personal information than any government without being democratically elected or accountable to the public (Prainsack 2017). This indicates once again that digital platforms and smart health services should not be seen as neutral tools. Rather, they incorporate certain values, norms, and ideologies, most importantly the capitalist ideology, as notions such as “algorithmic ideology” (Mager 2012, 2014), “informational capitalism” (Fuchs 2010), “cognitive capitalism” (Pasquinelli 2009), or “surveillance capitalism” (Zuboff 2019) suggest. In order to make medical care intelligent with the help of machine learning, they collate not only people’s behavior and concerns but also social and biochemical traces, environmental information, and archives of medical decision-making to improve online medical services and personal health management. Wilbanks and Topol (2018: 346) warn of the impact of the ongoing privatization of health data:

“If undisclosed algorithmic decision-making starts to incorporate health data, the ability of black-box calculations to accentuate pre-existing biases in society could greatly increase. Crucially, if the citizens being profiled are not given their data and allowed to share the information with others, they will not know about incorrect or discriminatory health actions—much less be able to challenge them. And most researchers won’t have access to such health data either, or to the insights gleaned from them.”

Moreover, recent research illustrates that many mobile health apps not only provide no or little privacy protection, they also move data back and forth between third parties, including industries such as insurance, banking, advertising, pharma, and even law enforcement (Huckvale et al. 2015). The platform Fitbit is a good example of such a business model: it sells the fitness data generated by its users back to those users and to other interested market stakeholders, though to the latter in a supposedly unidentifiable format. In this context, data ownership becomes a contested entity and has already evoked spoofing and hoaxing practices, as presented on the website Unfit Bits. In another context—the legal context—tracking data have already been used as evidence against their owners in court, as reported by several authors; the personal identifiable information was made available to prosecutors by the companies (Olson 2014; Alba 2016; Crawford et al. 2015). Industry and the security/law enforcement complex are constantly testing the legal limits of privacy invasion (van Diijk 2014). Then again, advocates of “personal data stores” (Pentland 2015) that enable users of e.g., social media to make a profit from their data while becoming “smarter” emphasize the potential value creation, which is to our understanding in line with the neoliberal ideal of self-governance. Microsoft Research launched the project Bali, a sort of personal database that should put the users in control of all the data collected about them (Gurevich et al 2014). Platforms like this use diverse notions of ownership and do not always provide clear legal frameworks for the services they provide. Ownership of personal data is hence increasingly being discussed at national and international levels (Kostkova et al. 2016). It is somewhat remarkable, however, that even the German chancellor Angela Markel—representing a supposedly data protection-friendly country—seems to buy into Silicon Valley’s rhetoric of “privacy as anti-innovation,” as Cohen (2013, 2014) described it in the context of a larger techno-political climate that she labels the “surveillance-innovation complex.” In this paradigm, user surveillance is framed as a necessary prerequisite for innovation, while privacy is seen as “antiprogressive, overly costly, and inimical to the welfare of the body politic” (Cohen 2013: 1904). Accordingly, Angela Merkel discounted data sovereignty as merely “philosophical” and limiting progress and innovation at the World Economic Forum in Davos in 2018: “Europeans have not yet made a real decision on how they want to deal with data. There is a great danger that we are too slow and will find...
ourselves overturned by events, so to speak, while we hold our philosophical debates about data sovereignty.” This quote further illustrates the tough negotiations between, and intrinsic tensions within, EU member states when it comes to their position on data protection standards and digital innovation, as Mager (2017) has shown in the development history of the General Data Protection Regulation (GDPR). Moreover, the GDPR was (and is) mainly criticized by small and medium enterprise (SME) actors, many of whom are now also in the business of “lively data” (Lupton 2016c) and therefore part of the rising global knowledge economy. They try to perform what Lyon (2003) has themed “social sorting,” a more and more granular clustering and categorization of social data, but they cannot act as boldly as big tech corporations with their legal departments and global reach.

Moreover, many data analytics corporations have their foundations in pre-digital times and pursue traditional risk analysis or personal marketing, as described before. They are hence used to being the owners of the consumer data they collect. New rules on data and privacy protection, especially the ban on data flows to territories outside the jurisdiction, theoretically require them to change their strategies and reconfigure their infrastructures. In practice, however, data breaches often remain unpunished. Therefore, if governments and public health services do not change their strategies, these corporate actors will remain in the position of being the “data-rich” gatekeepers of particularly valuable social data. Real-time health analytics and health data brokerage—core applications of data science—affect the lives of data producers (or data sources) in many ways, as they feed into the large-scale collection of consumer data (Christl et al. 2017). The current situation, however, will not help to challenge or reverse the direction of power transfer from individuals and communities to industries. On the contrary, corporate hegemony in the health system is being fostered, as governmental or institutional needs are more and more outsourced to industry without adequate strategies for empowering consumers and patients alike to stay in control of their data. The recent scandal of a soldier revealing secret military infrastructure by sharing his jogging routes on the Strava heatmap platform” (Hern 2018) illustrates how the unintended sharing of sensitive personal data can even become a matter of national security.

5. Discussion

In this article, we brought together different approaches to record and theorize the ambiguous trajectories of data bodies in health-related contexts. We showed several drivers behind the increasing quantification of health and the body from surveillance medicine (Armstrong 1995) to contemporary tracking and control technologies to illustrate the transformation of body data into data bodies. Building on classical actor-network theory (Latour 2005; Law and Hassard 1999) and contributions from new materialism, especially those from the medical context (Lupton 2016, 2018), we defined data bodies as the socio-material coupling of body and data. Drawing on exemplary use cases and business models, we described how these data bodies oscillate between empowerment and social control.

As we have seen, whenever data bodies start to travel they become commodities that are “vulnerable to alteration, addition, merging, and loss as they travel” as Lyon (2005: 22) explains. Furthermore, data bodies rely on complex information infrastructures and legislative frameworks, all of which facilitate specific types of social sorting and answer to socioeconomic pressures, sometimes further encouraging the quest for accuracy and better (biometric) identification. Data bodies may return to individuals in the form of classifications and discrimination; data bodies voluntarily created for empowerment purposes may ultimately be turned against the individual. “Thus the disappearing body is made to reappear for management and administrative purposes by more or less the same technologies that helped it to vanish in the first place” (Lyon 2003: 18).

When considering these new dimensions of digital biosociality (Jordan and Pfarr 2014; Hagen 2010), we can see that the various implications range from empowerment of communities in health care and medical research on the one hand, to social control triggered by commercial infrastructures that commodify and report personal information in many forms on the other. The double notion of body data/data body creates the complementary perspective needed to understand both bodies and data from their creation to their diversification into digital eternities, and how they feed into the continuous stream of dataveillance (van Dijck 2014). While data bodies flow through diverse settings, change health relationships, and enforce new dimensions of “patient work” (Strauss et al 1982; Mathieu-Fritz and Guillot 2017), they are at the same time promoting a more autonomous approach to health management,
redesigning self-awareness processes, motivating individuals to optimize their well-being and insurance companies their risk management. They bring about new types of classification systems and sorting mechanisms of the social, whose co-shaping of individual and communal practices still have to be studied more closely. Instead of making lives easy or perfectly transparent, they add to the complexity of inscriptive entanglements. Metering individuals are adapting themselves to standards and biopolitical norms that guide communities, while at the same time co-creating new classification systems that enrich statistical populations and might lead to a more holistic understanding of health. In this process, notions of trust, responsibility, accountability, and faith in data-handling institutions play an important role in how body data and data bodies can unfold their socio-technical potential.

**Conclusion: Tapping into the communal potential of digital health data**

The double focus on body data/data body allows us to understand both the mutual transformations and their potentials for creating new communal agency in this age of increasing privatization and individualization of health-related data. In this concluding chapter, we discuss aspects of ownership, care, and control of digital data bodies and how both society and individuals may cope with them in the future. More specifically, we suggest three options for action resulting from our analysis: (1) developing solidarity-based legislation, governance models, and institutions, (2) creating and fostering open socio-technical infrastructures instead of black-boxed technologies, (3) building capacities for new skills and literacy grounded in collective expertise. This can help us to broaden the perspective and challenge the engineering-centric logic underlying much of the smart service systems literature that currently informs policy debates and public spending, as argued earlier (Maglio 2015). Our STS-informed approach can further help to resituate narrow and individualistic value-creation processes suggested by smart service policies (e.g., technology-focused solutions, simplistic customer-centric business models, and consumer needs) to empower collective endeavors in the public health sector.

1. **Solidarity-based legislation, governance models, and institutions:** Digital health is blurring the lines between governments, health sector industries, and patient networks (Prainsack 2017). It is overthrowing the traditional roles and functions of stakeholders and infrastructures. Creating trustworthy and privacy-enabling environments for data bodies in the health sector consequently requires convincing and robust governance models and institutions (Sharon and Lucivero 2019). Building on our analysis, we suggest focusing on strong legislation that protects consumer data and sensitive information. With the recently introduced General Data Protection Regulation (GDPR), the European Union took a remarkable step toward better protecting users’ privacy and sensitive data. The GDPR mainly relies on individual informed consent procedures, with the disadvantage of making data protection an individual responsibility though. This is problematic as the autonomous individual who is supposed to take an informed decision does not always exist in practice, as Prainsack (2017) shows in the context of informed consent procedures in the medical field. Since informed consent procedures usually constitute “a social process rather than a situation of isolated decision making based on rational reasoning of individuals” (Prainsack 2017: 143), the author advocates for solidarity in this era of personalized medicine. A communal type of data sovereignty or governance of data bodies—a solidarity-based approach—should focus on collective ownership and control of data, place more emphasis on whether data use is in the public interest, and strengthen harm-mitigation instruments to reimburse individuals who are harmed by data use (Prainsack 2017; Prainsack and Buyx 2013, 2017). Given the profound sociopolitical challenges data bodies currently pose—and are expected to pose in the future—we therefore argue for reconsidering individual informed consent procedures and shifting the focus toward community-centric governmental and legislative actions grounded in broad societal debates. There are much bigger issues at stake than individual privacy: dataveillance brings discrimination and powerful new classifications into the governance of public health. With this article we show that following data bodies in their creation and digital biosociality could foster a more non-totalitarian and relational understanding of sovereignty (Rabinow and Rose 2006). To achieve solidarity-based governance models, legislation, and institutions, a broad societal debate is required where stakeholders, communities, and initiatives from broad societal fields should be invited to contribute and become involved. The experiences and expertise of self-tracking communities and patient networks are key here, but consulting other stakeholders—including medical experts, big data analysts, privacy advocates, etc.—is essential too.
in order to pave the way toward more communal types of data sovereignty and sustainable governance bodies that will enable a responsible framework for increasingly detailed data bodies, yet leave enough freedom for self-tracking communities and their desires and needs. Furthermore, solidarity-based institutions like public health care need to be strengthened (again), and new institutions like ethical boards may be introduced to tackle highly sensitive data bodies.\(^{12}\)

2) **Open socio-technical infrastructures instead of black-boxed technologies:** Even though a culture-pessimistic view predicts pervasive technological control with data science on the rise and technocratic perspectives gaining momentum in the context of machine learning coupled with surveillance instrumentation, we believe that we can still interfere in the practices of social sorting and build alternative infrastructures. Most tools in the connected age still mediate or stabilize social relations rather than fully controlling them (Latour 2002). We still have the option to focus on opening up and diversifying black-boxed procedures, methods, and classification systems. Engaged citizen sensing projects and several patient networks have successfully demonstrated how this can be done. Self-tracking communities (sometimes intentionally, sometimes unintentionally) and patient networks’ data-clearing initiatives have shown very innovative approaches to the ownership, care, and control of data bodies in socio-technical assemblages. Learning from self-tracking communities or data-sharing patient organizations that care about their data and their communities can enrich our understanding of how to build a more open and collaborative health sector. Depending on the respective public health system and other geopolitical configurations, the establishment and sustainability of health communities needs independent, more transparent socio-technical infrastructures and knowledge bases of human-data assemblages. Only then—in a safe space between small and big data bodies—can issues like privacy, data bias, and new forms of normalization and discrimination be critically discussed and strategies for counteraction be developed.

3) **New skills and literacies grounded in collective expertise:** Finally, with the focus on the collective or communal dimension of self-tracking, we have to ask how the construction and circulation of data bodies can contribute to the empowerment of individuals. Indeed, quantification and digital self-profiling can actually mean empowerment and freedom for people, as we have shown in this article. Patients can broaden their horizons, widen their access to care and expertise, and embed their own perspective (or instincts) in a larger body of knowledge. Shared services and quantification can further contribute to connecting to communities, breaking free from an individual story and entering a communal space with comparable accounts. In such a setting people can engage beyond being just statistical numbers in a population. To unlock the full potential of quantifying technologies and digital biosociality, new skills and literacies are needed. From emancipated self-surveillance communities we can learn that countering the trend toward “hyperindividualization” (Mau 2017) requires a broad set of skills, literacy, and knowledge ranging from technical know-how, community and collaboration management to medical sense-making. Those who actively build knowledge bases, train skills and data body literacy can contribute to individual and communal capacity building by both sharing their expertise and creating spaces for exchange and debate. The ability to scrutinize black-boxed technologies, understand and interpret data bodies in a comprehensive way, as well as organize concerns—or even modes of digital self-defense—will be crucial in the future, both for individuals and for society at large.

References:


PatientsLikeMe. Online: https://www.patientslikeme.com/ [11.06.2019].


Prainsack B. (2013): Let’s get real about virtual: online health is here to stay. Genetics research, 95 (4), 111-113.


