The Datafication of Health

Minna Ruckenstein1 and Natasha Dow Schüll2

1Consumer Society Research Centre, University of Helsinki, Helsinki 00014, Finland; email: minna.ruckenstein@helsinki.fi
2Department of Media, Culture, and Communication, New York University, New York, NY 10003; email: natasha.schull@nyu.edu

Abstract

Over the past decade, data-intensive logics and practices have come to affect domains of contemporary life ranging from marketing and policy making to entertainment and education; at every turn, there is evidence of “datafication” or the conversion of qualitative aspects of life into quantified data. The datafication of health unfolds on a number of different scales and registers, including data-driven medical research and public health infrastructures, clinical health care, and self-care practices. For the purposes of this review, we focus mainly on the latter two domains, examining how scholars in anthropology, sociology, science and technology studies, and media and communication studies have begun to explore the datafication of clinical and self-care practices. We identify the dominant themes and questions, methodological approaches, and analytical resources of this emerging literature, parsing these under three headings: datafied power, living with data, and data–human mediations.

Keywords
datafication, health care, digital health, self-tracking, wearable technology, big data, quantified self, biopolitics, algorithms, dataveillance

First published as a Review in Advance on July 28, 2017

The Annual Review of Anthropology is online at anthro.annualreviews.org

https://doi.org/10.1146/annurev-anthro-102116-041244

Copyright © 2017 Ruckenstein & Schüll. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

Click here to view this article’s online features:
• Download figures as PPT slides
• Navigate linked references
• Download citations
• Explore related articles
• Search keywords
INTRODUCTION

Over the past decade, the capacity to gather, store, and analyze individuals’ physiological, behavioral, and geolocational data has come to affect a wide array of everyday life domains, from policy making to policing, corporate marketing to electoral forecasting, entertainment to education, urban planning to epidemiology. So-called big-data fundamentalists promote the view that large data sets, properly mined for correlations and patterns, will render up previously elusive insights, predictions, and answers to long-standing challenges of individual and collective life, replacing the need for theory and science (e.g., Anderson 2008). Scholars in social science and the humanities take a more skeptical stance, emphasizing the cultural, political, economic, and rhetorical dimensions of the data paradigm shift, typically by focusing on particular cases of “datafication,” or the conversion of qualitative aspects of life into quantified data (Mayer-Schönberger & Cukier 2013, Van Dijck 2014).

The datafication of health unfolds on a number of different scales and registers, including data-driven medical research and public health infrastructures, such as biobanks and governmental databases; clinical health care, as in continuous patient monitoring, implantable biosensors, the use of the Internet for doctor-to-patient interaction, and personalized or “precision” medicine—practices collectively described as digital health, eHealth, mHealth, or Health 2.0; and self-care practices, as in the use of direct-to-consumer genetic and microbiomics testing websites, health-related peer-to-peer social media, and a vast array of wearable fitness and health devices and smartphone applications (apps). Although these different registers have never been neatly distinguished and have become increasingly blurred by a growing cross-traffic of data, for the purposes of this review, we largely bracket medical research, biobanks, and public health initiatives to focus on how scholars in anthropology, sociology, science and technology studies, and media and communication studies have begun to explore the datafication of clinical and self-care practices.

This emerging literature focuses almost exclusively on advanced industrialized jurisdictions—particularly North America, the United Kingdom, Australia, and Northern Europe—that face similar challenges: an aging population, rising rates of chronic disease, unsustainable health care costs (despite quite different health policy landscapes), and the retreat of social welfare. Market conditions have accelerated the blurring of clinical, research, and commercial domains around this “healthcare system crisis” (Sharon 2015, p. 295; Hogle 2016; Rich & Miah 2017; Schüll 2016a). The Global North also demonstrates a relatively broad embrace of the Internet and self-tracking technology by citizens; a cultural model of the ideal citizen as digitally literate and self-advocating; and a robust public debate around the ethical, legal, and social implications of big data.

When it comes to the datafication of health care and self-care, proponents in medicine, government, and technology typically emphasize its potential to prevent and mitigate the physical and financial burdens of “lifestyle diseases” such as obesity, diabetes, and cardiovascular disease—conditions that derive from daily behaviors of overeating, underexercising, and smoking—by shifting their management away from hospitals and doctors and into the hands of empowered patients (e.g., Goetz 2010, Topol 2012). An important starting place for social-scientific scholarship on the datafication of health is the recognition that this vision is mostly speculative, promissory, and, as yet, unrealized; what is needed, above all, are rigorous accounts of the actual reality of datafication as it takes shape in diverse practices and, quite often, twists in unforeseen directions.

1Here we follow the growing convention of treating “data” as a collective singular noun.

2We use “datafication” as a descriptive rather than a prescriptive concept; the term indexes a set of recent changes to the political-economic and experiential landscape—especially in affluent, Euro-American contexts—and is used by data enthusiasts and critical analysts alike.
We organize this review into three thematic clusters, each of which can be associated with certain research questions, methodologies, and theoretical commitments: datafied power, living with data, and data–human mediations. The clusters do not represent debates or disagreements so much as they represent parallel conversations that place weight on different themes, sites of inquiry, and analytical frameworks; considering these conversations alongside one another makes salient their strengths as well as their shortcomings and suggests ways in which their elements might be productively combined. Although some work on the datafication of health already moves in this direction, we argue that there is room for further cross-pollination. We conclude by pointing to a number of generative avenues for future research, urging scholars to investigate how datafication is unfolding across various digital divides (e.g., financial, technological, geographic), to experiment with applied forms of research and data activism, and to explore links to areas of datafication that are not explicitly related to health.

DATAFIED POWER

Scholars working in the interdisciplinary field of critical data studies have begun to address how data comes to “permeate and exert power on all manner of forms of life” in societies that are robustly digitally networked (Iliadis & Russo 2016, p. 2). Often starting from the idea of a divide between the “data rich” (e.g., governments, institutions, and commercial enterprises) and the “data poor” (individual citizens), these scholars draw attention to the “asymmetric relations between those who collect, store, and mine large quantities of data and those whom data collection targets” (Andrejevic 2014, p. 1673; see also Beer 2009, boyd & Crawford 2012, Lash 2007, Van Dijck 2014).

As the health-data ecosystem expands to include novel types of data and new methods for capturing, analyzing, and making use of this data, it is not only governmental agencies and health insurance companies that are data rich but also consumer technology companies. Writing from a philosophical-legal perspective, Nissenbaum & Patterson (2016) have examined how such companies monetize data that has been voluntarily collected by individuals (for their own purposes) by extracting and combining it with others’ data to draw population-wide correlations and inferences that hold value on the market of health data. Consumer electronics leaders such as Nike and Fitbit, for instance, share customer data with affiliated companies as well as marketers and insurers (Sharon 2016), and online platforms such as PatientsLikeMe and 23andMe sell health information to tech firms, pharmaceutical companies, and medical device makers (Van Dijck & Poell 2016). Patient data is valuable also to advertisers, insurance companies wishing to conduct cost-benefit analyses, credit-rating agencies wishing to create richer consumer profiles, and hospitals wishing to improve facility efficiency and save costs or to calculate predictive “health and frailty scores” for their patients (Hogle 2016, p. 385; Neff 2013).

In some cases, health websites and health-tracking devices offer consumers baits such as promotional gifts as a way to encourage their data “donations” (Lupton 2016b, p. 118) or promote the idea that consumers who share their data are contributing to “the public good” rather than to corporate profit making (Ajana 2017, Harris et al. 2013, Van Dijck & Poell 2016). More often, “data are surreptitiously extracted from data subjects, hijacked to serve agendas that benefit research and industry” (Iliadis & Russo 2016, p. 1; Crawford et al. 2015, Nissenbaum & Patterson 2016). In both cases, individuals perform unpaid and invisible “digital labor” in their role as data sources (Lupton 2016b, p. 118; Till 2014); their data streams become “biocapital” to harness and exploit (Rabinow & Rose 2006, p. 203; Kahn 2014). Selves are “sliced and diced into decontextualized parts, and bought and sold,” write anthropologists Nafus & Neff (2016, p. 62).

Nafus & Neff (2016) regard the commodification and exploitation of personal health data as just two dimensions of the “extensive process of biomedicalization” (p. 19; Clarke et al. 2003)
that has come with datafication, involving not only new modes of production, distribution, and consumption of medical information but also new forms of technoscientific objectification of bodies as they are “abstracted from real time, actual location, and social space” (Lock 1993, p. 371). Lupton (2012) points out that mobile health technologies such as wearable, networked glucose monitors fall along the same historical continuum as medical imaging technologies, rendering patient bodies observable to a clinical gaze (Foucault 1973); by converting everyday activities into data flows that can be monitored continually and at a distance, digital tracking tools expand the range of this gaze, offering health care professionals (and technology designers) remote access to previously inaccessible domains of individuals’ lives (see also Nafus & Neff 2016, pp. 18–19). The tools of mHealth, Lupton argues, exemplify not only the medical gaze but “surveillance society” more generally (Andrejevic & Burdon 2015; Armstrong 1995; Foucault 1977; Haggerty & Ericson 2000; Lyon 2002, 2014).

Some scholars of datafication argue that “surveillance” is too optically freighted and centrally organized a phenomenon to adequately characterize the networked, continuous tracking of digital information processing and algorithmic analysis; they propose instead the term “dataveillance” (Gandy 1993, Raley 2013, Ruppert 2011, Van Dijck 2014; see also Koopman 2014). Rather than originating from a singular source positioned “above,” dataveillance is distributed across multiple interested parties—in the case of health, including caregivers, insurance payers, pharmacies, data aggregator and analytics companies, and individuals who provide information (either wittingly or unwittingly). Another feature that distinguishes dataveillance from surveillance is its predictive telos; its aim is not to “see” a specific behavior so much as to continuously track for emergent patterns (McQuillan 2016, Van Dijck 2014). Just as providers use data to know in advance which patients are likely to be risky and expensive, health insurers use data to sort, rank, and differentially charge their customers for services. A data-analytics division of United Health Group, for instance, developed “a drug adherence index” to identify patients who were noncompliant with prescription drugs (Hogle 2016, p. 384; see also Morozov 2014, pp. 238–40; Nafus & Neff 2016, p. 163; Hull 2017).

Digital tracking tools and algorithms are used not only to detect and predict but also to shape and modify behavior (Beer 2009, Lash 2007, Mackenzie 2005), as we discuss at greater length in the Data–Human Mediations section below. The data users generate can be processed and fed back to them so that they can modulate their actions, forming a digitally mediated feedback loop that at once recalls and departs from panoptic forms of discipline (Foucault 1977). A case in point is the growing trend in corporate wellness programs (Hull 2017, Till 2017), many of which incentivize employees—with the prospect of bonuses or discounts on health insurance premiums, or with the threat of penalties—to engage in self-tracking activities such as exercise or cessation of smoking, thus generating data that employers or insurance carriers can see and analyze using opaque and proprietary algorithms (Christopher et al. 2015). “The advent of big data brings with it new and opaque regimes of population management, control, discrimination and exclusion,” write Kennedy et al. (2015, p. 1).

To characterize these regimes, many have turned to Foucault’s (1986, 1997) concept of biopolitics—a mode of power that seeks to understand, control, and regulate the “vital characteristics” of the population, as reflected in rates of birth, disease, productivity, and mortality collected through censuses and other “population metrics” (Rabinow & Rose 2006, p. 196; Ruppert 2011). “We are entering an era in which biopolitics have become increasingly digitized,” writes Lupton (2016b, p. 56). In an article on digital health, Ajana (2017, p. 2) similarly writes of “data-driven modes of biopower” (see also Depper & Howe 2017; Rich & Miah 2017, p. 93; Oxlund 2012).

Responding to such work, some scholars of datafication move in what might be called a post-biopolitical direction, suggesting that a biopolitical framework “cannot fully exhaust the new
modes of information surveillance, aggregation, and distribution in our midst” (Koopman 2014, p. 105). Schull (2016a) points out that data-tracking concerns the monitoring and modulation not only of biological life but also extracorporeal elements, including daily choices and rhythms, preferences, and tendencies. Ruckenstein (2016) observes that diverse fields of expertise—not only statistics and public health but also behavioral economics, social psychology, and human computer interaction—are interested in quantifying aspects of people’s lives and do so via a range of different models and expectations for how life is or should be lived. A biopolitical lens alone, both authors suggest, might obscure rather than illuminate distinctive features of datafication such as real-time and predictive analytics or the serpentine way in which personal data circulates through insurance databases, clinical care encounters, and day-to-day self-care practices.

For some, the contemporary datafied world evokes the sort of diffuse and continuous monitoring and modulation that Deleuze (1992) associated with “control society” (Andrejevic & Burdon 2015, Cheney-Lippold 2011). “Disembodied exhaust gives rise to a data-proxy, an abstracted figure created from the amalgamation of data traces” (Smith 2016, p. 110, emphasis in original; see also Millington 2016, Rich & Miah 2017, Williamson 2015). Drawing on Deleuze’s idea of the “dividual,” Greenfield (2016, p. 133) discusses the “pixelated person” of precision medicine as “a subject ever divided into finer granularity, but also whose partial datasets can be joined with others.” Other scholars prefer to regard the datafication of health as a form of neoliberal subjectification (Foucault 1991), emphasizing how tracking tools accelerate the withdrawal of the welfare state from citizens’ lives, turning health care into self-care (Ajana 2017; Depper & Howe 2017; Fotopoulou & O’Riordan 2016; Lupton 2013a,b; Mort et al. 2009; Rich & Miah 2014, 2017; Schull 2016a,b). Subjecting oneself to data tracking becomes an avenue for performing “healthy citizenship” and enacting cultural values of entrepreneurial, autonomous behavior. “In the discourse of the digitally engaged patient,” Lupton (2013a, p. 261) observes, “empowerment becomes a set of obligations” (see also Oudshoorn 2011, Schull 2016a, Sharon 2017). Taking a different perspective, Sharon (2015, p. 295) suggests that this mode of scholarly criticism often “obscures the many ways in which individuals engage with healthy citizenship discourse that are not governed by principles of autonomous choice and that do not corroborate fears of normalization and discipline.” As we elaborate in the next section, these engagements are marked by ambivalence over agency, relief in outsourcing responsibility, and an embrace of digital care.

Scholars who attend to the power dynamics of datafication have been faulted for their heavy focus on the oppressive, normalizing, and exploitative forces of datafication and their lack of attention to cases of noncompliance, appropriation, or existential possibility. It is important “to understand data and datafication not only in terms of power and domination, but also in terms of agency,” write Kennedy et al. (2015, p. 6; see also Couldry & Powell 2014). For instance, “people do not always comply with suggested or imposed regimes of health care” (Sharon 2015, p. 296), instead mobilizing data in creative and even pioneering ways (Rapp 2016). What is needed to grasp datafied power in all its richness is, in part, a loosening of theoretical commitments—a willingness to suspend ready-made analytic frames to explore the wide range of agencies and aims that are in play and at stake. As we consider below, ethnographic methods and practice-based analysis can deepen and nuance understandings of datafication by revealing how data and its technologies are taken up, enacted, and sometimes repurposed.

**LIVING WITH DATA**

Even as heated public debate unfolds over how data tracking by governments, medical institutions, and corporations might undermine personal identity, liberty, and privacy, individuals have embraced data technologies in ways that not only benefit powerful stakeholders in the larger political...
An economy of datafication but also serve as “technologies of the self” (Foucault 1988). This “new intimacy of surveillance,” as the cultural and linguistic anthropologist Berson (2015, p. 40) characterizes it, has been the focus of a growing set of ethnographic inquiries into the social, narrative, and affective dimensions of actual data practices and experience.

A critical ethnographic site for examining social dimensions has been the Quantified Self (QS), an international collective whose members seek “self-knowledge through numbers,” as its website tagline reads. Since its 2008 inception, the group has facilitated online forums and live meet-ups, constituting itself as a learning community whose members are linked by their common interest in insights gained through quantification and data analysis (Barta & Neff 2016, Dudhwala 2017, Greenfeld 2016, Jethani 2015, Nafus & Neff 2016, Sharon & Zandbergen 2016). “QS is one of the few places where the question of why data matters is asked in ways that go beyond advertising or controlling the behaviors of others,” write Nafus & Sherman (2014, p. 1788). In an online ethnography of QS videos, Smith & Vonthethoff (2017, p. 12) explore how members “narrate personal experiences and stories in a public forum via the ‘companion’ medium of their data.” Shared data offers a common language that people can relate to, becoming “a medium of connecting with others by offering a raw glimpse into one’s intimate private life” (Sharon 2017, p. 20). We might think of this form of open-ended communication (Lomborg & Frandsen 2016) as “datasociality,” updating Rabinow’s (1992) idea of biosociality as the production of social relationships and identity categories out of genetic or biological conditions.

Others have described datasociality through examinations of apps and websites such as MyFitnessPal and CureTogether, which allow members to share knowledge in ways that can bypass, override, or exceed professional expertise (Maturo & Setiffi 2016). Consumer genomics and ancestry testing sites such as 23andMe and FamilyTreeDNA can alter understandings of kinship ties and ethnicity (Bolnick et al. 2007, Lucivero & Prainsack 2015, Nelson 2016). Kragh-Furbo et al. (2016, p. 6) suggest that we regard such sites as “social spaces in which power laden practices of knowing, prediction, making promises, sensing, and caring might be negotiated differently.” These negotiations are, in Sharon’s (2017, p. 93) terms, “enactments of solidarity.”

Venturing beyond questions of sociality to consider the narrative dimensions of self-quantification, ethnographers have observed that engagements with personal data and its technologies enable individuals to make meaning of their life experiences and relate to their bodies and selves. Genetic testing sites, for instance, can serve as portals of access to raw genetic code that consumers can probe for health-related details with open-source tools (Ruckenstein 2017) or use to construct “autobiologies” (Harris et al. 2014) and “emergent identities and ways of knowing” that are not necessarily essentializing or normative (Fishman & McGowan 2014, p. 39).

Whereas a biomedicalization framework tends to place a negative emphasis on the reductive, fragmenting, decontextualizing effects of quantification on selves, ethnographic studies show that self-quantification “rarely produces a definitive truth, a one-to-one representation of one’s life or one’s identity” (Sharon 2017); instead, it involves a “situated objectivity” (Pantzar & Ruckenstein 2017) in which certain prior experiences, understandings, and shared expectations come to matter. Sherman (2016) has discussed self-tracking as an aesthetic practice in which bits of the self, extracted and abstracted, become material for differently seeing and experiencing the self. Ruckenstein’s (2014, p. 80) fieldwork shows that looking at personal data charts and visualizations can trigger critical reflection and raise new questions to pursue; the data does not displace or freeze but rather enhances and enlivens self-narratives. Schull’s (2016b) interviews reveal that extended time-series analysis of self-data frees trackers from a sense of fixed, essential identity (see also Day & Lury 2016). Nafus & Neff (2016, p. 25) describe data as a kind of “transducer” that preserves only some qualities of the thing being measured such that “there is much room for people to maneuver in the imperfect translation.” For instance, it might be easier to cope with
a bladder problem when it is objectified in the form of a spreadsheet tallying bathroom visits, or a bipolar disorder when it is measured with scales, charts, and numbers (Martin 2007). Data renders aspects of a “somewhat inaccessible world of feelings and problems more tangible and comparable” (Sharon & Zandbergen 2016, p. 11).

Ethnographic research also reveals that people do not always use tracking tools or engage with data in the ways that technology designers, website developers, or medical institutions imagine or intend. This is especially true for ordinary users (Didžioikaite et al. 2017, Gilmore 2016) who may abandon, ignore, or forget about tracking devices and apps or use them only sporadically (Lazar et al. 2015). Some produce incomplete or uneven data owing to irregular device use while others resist the data-harvesting interests of technology companies by switching between devices and services and disrupting the integrity of data streams; in this way they “partially yet significantly escape the frames created by the biopolitics of the health technology industry” (Nafus & Sherman 2014, p. 1784), enacting a kind of “soft resistance.” They may “tinker” with devices, using them as toys for entertainment purposes (Ellerbrok 2011) or for recording life events and transitions without any intention to modify behavior (Elsden et al. 2016). Confronted with self-data suggesting that healthier day-to-day choices are in order, some might conclude that a lifestyle change would require too much effort or might revise their understanding of what counts as a healthy and balanced life. In some cases, the data revalorizes underappreciated and neglected aspects of daily lives (Ruckenstein 2014, Sharon 2015). Fiore-Gartland & Neff (2015, p. 1470) propose the notion of “data valences” to underline differences in expectations and values assigned to data by different data stakeholders (e.g., designers, doctors, patients, and self-trackers).

In the clinical context, self-tracking can balance or adjust the power dynamics of the doctor–patient relationship by giving rise to forms of therapeutic alliance in which patients and providers are both seen as experts (Nafus & Neff 2016, p. 140; Neff 2013). Self-generated data can mediate between subjective experiences of physical and mental health and the medical profession’s protocols and diagnostic categorizations (Ruckenstein 2015). In some cases, patients use the data to second-guess or subvert medical expertise and intervention; those with chronic conditions such as diabetes might start self-tracking when pushed by doctors and later refuse to share the data (Piras & Miele 2017). Greenfield (2016, p. 127) suggests that we think of such data practices as operating in a paraclinical register, “working alongside and even in opposition to formal medical practice.” Self-tracking can produce a kind of empowerment “in the wild” that differs significantly from the normative empowerment that digital health promoters have in mind (Sharon 2017, p. 109), as when patients hack or dismantle tracking devices or set idiosyncratic tracking parameters that resist population-wide health norms (Nafus & Neff 2016, p. 20; Nafus & Sherman 2014; Sharon 2017).

Yet, there is a curious resonance between the vision of empowered, resisting individuals that many ethnographers of self-tracking celebrate and the rhetoric of consumer empowerment found in discourses of digital health (Schull 2017, Sharon 2017). An emerging group of ethnographers find that people have a far more ambivalent relationship with self-tracking technology and data—and, more generally, to the task of responsible self-management—than accounts of user resistance might suggest. In a forthcoming monograph, Schull observes that most self-trackers engage data technology with ambivalence, admitting a wish both to take charge of themselves and to delegate that task, burdensome and confounding as it is; they speak of feeling cared for by the automated interventions of their devices and released from hard-to-meet demands for self-regulation (see also Schült 2017).

This ambivalence is reflected in the volatile range of affective orientations that people have toward the tracking of self-data—from hope to disappointment, pleasure to frustration, control to obsessiveness. People can love their data assemblages, even transforming them into valued possessions, or hate them. The relationship between individuals and their data can trigger experiences
of “doubt, guilt, fear, dismay, disappointment and hesitation as well as joy, enthusiasm, and pride” (Salmela-Leppänen et al. 2018, p. 1; see also Lupton 2017, Pantzar & Ruckenstein 2017). When a pedometer informs its user that the daily step count has been reached, one can feel emotionally elevated: The data offer validation of personal worth. Yet when a device repeatedly shows “bad numbers” despite best efforts, one can feel guilt or emotional deflation. Some admit to becoming addicted or enslaved to tracking technology and routines, pacing in their kitchens late at night to reach their preset daily goals; in such cases, the original aim to become healthier is “displaced” by an aim to maintain certain metrics (Dudhwala 2017, pp. 121–22).

The inductive orientation and fieldwork-based methodologies of anthropology, proceeding from observation to analysis and privileging particular cases over general frameworks, have much to contribute to understandings of datafication—most importantly, a richer sense of its agentic possibilities. Ethnographies of self-tracking, however, can focus too heavily on sociality, self-making, and experiential aspects of datafication, paying insufficient attention to how digital media shapes, steers, and constrains human agency, topics to which we now turn.

**DATA–HUMAN MEDIATIONS**

Scholars in science and technology studies, postphenomenology, and media studies argue that it is impossible to grasp the workings of datafication if one considers only larger power dynamics or the subjective experiences of data-generating individuals; one must also take into account the nonhuman elements that mediate these dynamics and experiences—device parameters and affordances, analytical algorithms, data infrastructure, and data itself, as well as the processes and practices around them. In their meditating capacity, these elements act with a kind of agency (Latour 2005, Salmela-Leppänen et al. 2018), “liveliness” (Lupton 2016b), or “performativity” (Kitchin 2014, Mackenzie 2005), guiding, formatting, or altering the course of a given tracked phenomenon according to their own classificatory and procedural logics. Below, we examine how a number of such intermediaries “structure and shape possibilities for action” (Williamson 2015, p. 141).

In their ethnographic study of hypoglycemia, Mol & Law (2004, p. 48) write about “the use of measurement machines to train inner sensitivity” to blood sugar levels; they call this sensitivity “intro-sensing.” Sensory ethnographers Pink & Fors (2017, p. 2) observe that the “digital materiality” of self-tracking technologies intimately mediates “people’s tacit ways of being in the world.” Nafus & Neff (2016, p. 75) describe data as a “prosthetic of feeling [that can] help us sense our bodies or the world around us.” Berson (2015, p. 34) regards self-data as “encodings of palpable impressions of change, difference, or variation” that, no matter how mediated, “exist by virtue of their availability to our senses.” His book shows how contemporary bodily experience is increasingly folded into data, and how data—as a particular kind of abstraction of experience—increasingly shapes experience and mediates human agency. Williamson (2015, p. 147) finds that “health-tracking data act as a kind of active, algorithmic skin that not only sheathes but animates and orders the body.” Not all scholars take a positive view of how quantified data mediates human experience; Smith & Vonhethoff (2017, p. 18), for instance, find it troubling that “bodily intuition is being outsourced to, if not displaced by, the medium of un-bodied data.”

This negative view carries over to the growing literature on algorithms, another nonhuman force at work in digital health and self-tracking. Those writing in the mode of critical data studies emphasize “algorithmic power” (Beer 2009, Lash 2007, Mackenzie 2005): Algorithms can “discriminate”; they can “pose a threat to the autonomy of data subjects” (Mittelstadt et al. 2016, p. 9); “operations, decisions and choices previously left to humans are increasingly delegated to algorithms, which may advise, if not decide, about how data should be interpreted.
and what actions should be taken as a result” (p. 9). Such analyses pit the automated procedures of algorithms (Gillespie 2014) against human autonomy.

Others choose to focus less on the power of algorithms than on how they are invested with social norms, cultural values, and governance agendas (Cheney-Lippold 2011, Gillespie 2014, Lyon 2002). In a study of health-tracking devices used in school-age physical education, Williamson (2015, p. 136) encourages analysts “to acknowledge that algorithms are both socially produced and socially productive, rather than to imply that they act deterministically” (see also Bucher 2017).

Hogle (2016) describes how algorithms, programmed with assumptions about social difference, perform a digital form of “social sorting” that both introduces new categories of people and illnesses and reinforces long-standing categories (p. 388). Self-diagnosis apps, for example, combine appeals to algorithmic authority with encouragement to seek traditional medical authority, should users want a formal diagnosis (Lupton & Jutel 2015). Self-tracking devices can operate as “epistemology engines” (Van den Eede 2015, p. 151) that bring machinic agency to bear on human ways of defining, categorizing, and knowing life.

As social expectations of normality and health become embedded in tracking devices’ target numbers, presentation of scores, and gamified incentives (Depper & Howe 2017, Whitson 2013), a “numerical ontology” comes to suffuse everyday practices and “the ways in which people relate to their own bodies” (Oxlund 2012, p. 53; see also Jethani 2015, p. 40). Most digital pedometers, for instance, are programmed with a daily target of 10,000 steps, no matter the circumstances of the user (Watson 2014). Although online weight-loss services calculate a “personalized” calorie intake recommended for each user, these are typically based on averages for large populations (e.g., women 20–25 years old) rather than genuinely tailored to each person’s physiology and metabolism (Niva 2017). Sexual activity tracking apps are designed with gendered assumptions: Those marketed to women tend to be framed in terms of reproductive risks and sexually transmitted diseases, whereas those for men typically focus on sexual performance, allowing users to record number of thrusts to orgasm, audio decibels, and the like (Eveleth 2014, Lupton 2015).

To understand how norms get encoded into data technology, some investigators turn their attention to designers and the design process, examining how these actors imagine potential users and work to modify and intervene in their behavior through hardware and software. Wearable computing engineers imagine the human body as “continually emitting signs, albeit in forms inaccessible to the self that might act to maintain it,” write anthropologists Viseu & Suchman (2010, p. 175). In a satirical piece meant to “provoke reflection about values and politics of design in persuasive computing,” Purpura et al. (2011) invent a fictional tracking system, Fit4Life, in which a suite of cajoling messages delivered through an earpiece encourage individuals to practice healthy behaviors. Drawing on research conducted among technology developers and marketers of personal health technology, Schull (2016a) considers how they “design self-care” into their products in the form of motivational feedback loops and “micronudges” that reinforce certain behaviors and discourage others. She detects a thermostat-like logic to these products, which do not merely monitor but actively regulate users via automated prompts such as taps, buzzes, and vibrations that serve to reduce uncertainty over which day-to-day choices to pursue. Berg (2017, p. 6) similarly observes that designers of smart rings approach users “as vulnerable beings in need of assistance, advice, and actionable guidance,” working from the assumption that human senses alone cannot handle the vagaries and intensities of daily lives and must be digitally supported. Posture-correcting wearables, Millington (2016, p. 414) writes, are likewise “imagined as active agents—as coaches and trainers” who can “call users to attention” when slouching is detected (see also Smith & Vonthethoff 2017, p. 17; Salmela-Leppänen et al. 2018).

Scholars have explored not only the design and performativity of data technology hardware (devices, sensors), software (user interface, operational and analytical algorithms), and data itself
(in the form of charts, numbers, and sensory nudges), but also the data infrastructures (labs, data centers, server and cloud storage, networks) that organize how databases store and circulate data in ways that enable certain perspectives and responses and disable others (Beer 2009; Bowker 2014; Dourish 2016, p. 8; Hogle 2016, p. 375). Data sets “are not, and can never be, neutral and theory-free repositories of information waiting to give up their secrets” (Crawford et al. 2014, p. 1668; Andrejevic 2014); the data they contain is never raw and objective but always made, inferred, and interpreted through infrastructure design (Boellstorff 2015, Gitelman 2013, Gregory & Bowker 2016). Without such design work, data might die, or get stuck in the sense of having no purpose, meaning, or designated place (Nafus 2014). An anthropologist and Intel executive comments, “The people who are coding data, writing algorithms, creating recommendation engines, are determining what kind of relationships there should be between data sets and making judgments about what data should speak to what other data” (Bell 2015, p. 16).

Health data streams can become part of a multitude of different agendas, each wanting to assert its particular script for coding, protecting, and modifying health. Questions and contestations can arise over who defines health and based on what kind of data processing. Examining the case of personal genomics, Gregory & Bowker (2016) emphasize how the political, economic, and social effects and possibilities of data are determined by the plethora of decisions and transformations involved in the design of its platforms. In a study of clinical data, Neff (2013) underlines the need to translate the data across the social worlds of patients and health providers, arguing for the “social interoperability” of health data.

In their attempts to theorize the dynamic entanglement of nonhuman and human, a number of scholars of datafication have embraced the analytical framework of the assemblage. “Assemblage is a concept that helps capture the multitude of ways that already-composed data structures inflect and interact with society, its organization and functioning, and the resulting impact on individuals’ daily lives” (Iliadis & Russo 2016, p. 1; Kitchin 2014). “We may begin to think about our data assemblages as companion species that have a life of their own that is beyond our complete control,” writes Lupton (2016a, p. 3). In some instances, the notion of the assemblage is used in an enumerative, all-encompassing fashion that dilutes its analytical power, as in formulations that “may include systems of thought, forms of knowledge, finance, political economy, governmentalties and legalities, materialities and infrastructures, practices, organizations and institutions, subjectivities and communities, places, and the marketplace where data are constituted” (Iliadis & Russo 2016, p. 3). A more analytically productive, pointed use of the concept does not simply enumerate the various elements that an assemblage comprises but, rather, homes in on how data moves through it: where it flows, where it finds impasses, how algorithms act on it along the way (Dourish 2016, p. 2). An exemplary use of the assemblage heuristic in the health and datafication field is found in Hogle (2016), who writes that, “more than the listing of components or charting of relations,” the term “indicates that such groupings are contingent and emergent, and they extend beyond a single domain of ‘health care’ or ‘data intensive science’” (p. 375).

**CONCLUSIONS: WIDENING THE FRAME**

In the recent upswell of scholarship around the datafication of health care and self-care, social scientists have brought a great many concepts and analytical frameworks to bear on a great many cases and concerns; they have spent less time identifying the possible points of tension or internal contradiction among them, or reflecting on how they might be combined in complementary and generative ways. We offer this review as a starting point for such a dialogue. In our final set of comments we turn from the existing literature to a number of less-traveled research paths that promise productive new directions for scholars exploring the datafication of health.
Across Digital Divides

As stated at the outset, scholarship pertaining to the datafication of health care and self-care focuses almost exclusively on advanced industrialized regions that demonstrate rising rates of chronic disease and health care costs as well as high levels of Internet use and uptake of digital self-tracking tools by citizens. Within this context, scholars should focus more attention on how lower-income and culturally marginalized individuals and communities have been affected by data technology, both in clinical settings and in their use of wearable devices and tracking apps. Although not well represented at QS meetings or within the consumer market for self-tracking tools, these groups are overrepresented in the case of conditions such as obesity and diabetes and are prime targets of digital health initiatives, health insurance wellness programs, and hospitals wishing to engage them in self-tracking regimens.

Those who study power asymmetries, discrimination, and threats to privacy are analytically well equipped to explore the datafication of health care beyond financial, educational, and cultural divides—yet some writing in this area risk confounding such divides when they characterize even the most financially affluent, technologically savvy, and culturally dominant among us as “data poor,” with powerful companies and governments occupying the position of the “data rich.” Although this is an analytically useful divide to consider, it blurs the fact that some of us are “poorer” than others with respect to the control we exercise over our data; it also misses the fact that a lack of data (e.g., medical records) can make subjects as vulnerable in relation to power as when their data is exploited (Iliadis & Russo 2016, p. 1).

As for the literature on datafied self-care, here nearly all the focus is on wealthy, educated, cosmopolitan citizens. One exception is a politically reflexive study by Peake (2015), who helped design self-tracking systems for African American, Native American, and migrant laborers suffering from tinnitus. We can attribute the lack of such studies in part to the simple fact that the emerging digital health market does not cater to poor or otherwise marginalized groups—and yet there is important empirical research to undertake on how these groups respond to the self-tracking behaviors in which they are pushed to engage by insurance companies, hospitals, and wellness initiatives.

In developing countries, where health is less a question of individual lifestyle than a question of access to basic health care and the challenges of communicable diseases, and where infrastructural affordances and digital literacy are often lacking (Arora 2016, Chib 2013, Chigona et al. 2013, Taylor & Broeders 2015), the dearth of research on the datafication of health care and self-care is even starker. Although multiple studies of “telecare” have been conducted in areas with high levels of cell phone penetration (al Dahdah et al. 2015, Duclos 2016, Hampshire et al. 2017, Miscione 2007), the focus is typically on connectivity and communication rather than on datafication; themes of privacy, self-fashioning via tracked data, or the ways that data technology enables or constrains human experience are mostly absent. 3

As health is increasingly datafied in the Global South, whether through mobile phones used as connective devices, interactive decision-support algorithms and diagnostic testing apps in smartphones, or wearables with real-time monitoring capabilities, scholars of datafication will need to widen their geographic frames. They should take such developments not merely as occasions to produce standalone case studies of the adaptation or translation of data technology and infrastructure into new settings by a different set of actors but as an opportunity to consider how these

---

1 Even studies on big-data initiatives in public health, epidemiology, or medical research in the Global South, Arora (2016, p. 1684) notes, tend to sideline critical perspectives such as those articulated in the West around constitutional rights; she attributes this to the humanitarian context as well as a neoliberal logic of inclusive capitalism.
settings and actors reshape, repurpose, or even reengineer that technology, offering insight into
the datafication of health more broadly.

Applied Insights and Data Activism
Recognizing that the datafication of the traditional health ecosystem is generating new power
asymmetries and disrupting traditional regulatory and ethical mechanisms, some scholars have
embarked on applied research projects, often collaborating outside their academic spheres. These
citizen-centered initiatives⁴ appeal to patient organizations, savvy self-trackers, and ordinary peo-
ple to address issues such as the tension between data openness and data ownership; asymmetries of
data usage and distribution; the inadequacy of current informed consent and privacy protections;
and the need to reappropriate and rearticulate concepts such as “sharing” and “the public good”
that have been co-opted by technology companies seeking free access to their users’ data. As Neff
(2013, p. 121) observes, “If the conversation in health technology innovation does not address the
questions of data for whom, when and why, then it will be a failure of social justice and abuse of
the trust that people have placed in the institutions of health care.”

A related strain of scholarship that might be characterized as “data activism” explores how the
capacities of data technology might be harnessed to promote social justice, equality, new forms
of agency, political participation, and collective action—and to challenge accepted norms and
ideological projects (Baack 2015, Delfanti 2013, Greenfield 2016, Kennedy et al. 2015, Milan &
vander Velden 2016, Pybus et al. 2015, Ruckenstein & Pantzar 2015). Individual self-tracking data,
for instance, can have social and political potential when it is pooled to identify health inequalities,
collective environmental exposure, or disparities in quality of life (Gabrys 2014). Such data might
demonstrate that physiological stress is tied to everyday rhythms defined by societal norms and
pressures rather than by individual predilections (Pantzar et al. 2017), or they might highlight
the physiological and emotional costs of particular work schedules for caregivers (Mehta & Nafus
2016).

Working on a personal register, open-source activist and media artist Salvatore Iaconesi set up a
website featuring information related to his brain tumor alongside an open call for crowd-sourced
responses to his predicament; by creating a public space in which to experience his illness and
healing, he resisted being reduced to the category of isolated cancer patient constituted by a set of
medical data and expert decisions (Delfanti & Iaconesi 2016). Examples such as Iaconesi’s demon-
strate the possibility of linking personal and clinical health data to broader ethical, technological,
and communicative processes and forces and of reorienting those processes and forces toward
the construction of collective spaces for calling into question the social imbalances of health and
medical regimes. Health, considered from the standpoint of data activism, is a societal rather than
individual issue, its meaning as political as it is existential.

Beyond Health
The literature we have reviewed here illuminates the movement of data across commercial, clinical,
and everyday health contexts, such that these contexts become blurred. Future studies could go
farther, exploring how the datafication of health is blurring with the datafication of other aspects

⁴Among these initiatives are Open mHealth, DIYGenetics, Genetic Alliance, Open Humans, Health Data Exploration Project,
QS Public Health Symposium, Atlas of Caregiving, Healthbank, MIDATA.coop, and MyData. For a discussion of participant-
driven research initiatives that involve new modes of data gathering and sharing, see McGowan et al. (2017).
of life. Scholars in diverse fields of social science and humanities have begun to take stock of datafication in such domains as communication and social media, urban life and the civic sphere, memory and self-documentation, education, finance, labor, and intimate relationships. Increased dialogue with such work could help to identify what is and is not distinctive about the datafication of health.

A scholar might, for instance, consider personalized medicine alongside personalized advertising; reflect on how health-monitoring technologies mimic video games and on strategies of gamification such as rewards, bonuses, and badges (Whitson 2013); ask why online patient portals are designed to mirror the self-budgeting affordances of banking websites; or explore how workplace wellness programs—or even personal time-management regimens—reframe health as fitness for work and, conversely, productivity as health (Gregg 2018, Hull & Pasquale 2017, Moore & Robinson 2016, O’Neill 2017, Potts 2010, Till 2017). Following the datafication of health beyond the obvious boundaries of clinical and self-care practices will lead to a more complex and nuanced picture of how health is being redefined, enacted, and experienced today.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Baack S. 2015. Datafication and empowerment: how the open data movement re-articulates notions of democracy, participation, and journalism. Big Data Soc. 2: [https://doi.org/10.1177/2053951715594634]
Bell G. 2015. The secret life of big data. See Boellstorff & Maurer 2015, pp. 7–26
Chib A. 2013. The promise and peril of mHealth in developing countries. *Mobile Media Commun.* 1:69–75


Gregory J, Bowker GC. 2016. The data citizen, the quantified self, and personal genomics. See Nafus 2016, pp. 211–22


Nafus D, Sherman J. 2014. This one does not go up to 11: The Quantified Self movement as an alternative big data practice. *Int. J. Commun.* 8:1784–94


Contents

Perspectives

Recovering the Body
Margaret Lock .......................................................... 1

Archaeology

Rock Art and Ontology
Andrew Meirion Jones .................................................. 167

Collective Action Theory and the Dynamics of Complex Societies
Elizabeth DeMarrais and Timothy Earle .............................. 183

Archaeologies of the Contemporary World
Rodney Harrison and Esther Breitboff ................................ 203

The Archaeological Study of Sacrifice
Glenn M. Schwartz .......................................................... 223

Archaeology and Human–Animal Relations: Thinking Through
Anthropocentrism
Brian Boyd ......................................................................... 299

Social Network Analysis in Archaeology
Barbara J. Mills .................................................................. 379

Biological Anthropology

Late Australopiths and the Emergence of Homo
Darryl J. de Ruiter, S.E. Churchill, J. Hawks, and L.R. Berger .............. 99

Primate Positional Behavior Development and Evolution
Michelle Bezanson .............................................................. 279

The Monkeying of the Americas: Primate Biogeography in the
Neotropics
Jessica Lynch Alfaro ........................................................... 317

Brain Plasticity and Human Evolution
Chet C. Sherwood and Aida Gómez-Robles .................................. 399
Anthropology of Language and Communicative Practices

Language and the Newness of Media
Ilana Gershon .................................................. 15

Personal Narratives and Self-Transformation in Postindustrial Societies
Cynthia Dickel Dunn ........................................... 65

Human–Animal Communication
Don Kulick ........................................................... 357

Sociocultural Anthropology

Consuming DNA: The Good Citizen in the Age of Precision Medicine
Sandra Soo-Jin Lee .................................................. 33

Precarious Placemaking
Melinda Hinkson .................................................. 49

Marriage and Migration
Caroline B. Brettell .................................................. 81

Fluid Drugs: Revisiting the Anthropology of Pharmaceuticals
Anita Hardon and Emilia Sanabria .................................. 117

Humans and Animals in Northern Regions
David G. Anderson .................................................. 133

What Does Catastrophe Reveal for Whom? The Anthropology of Crises and Disasters at the Onset of the Anthropocene
Roberto E. Barrios .................................................. 151

A Bundle of Relations: Collections, Collecting, and Communities
Joshua A. Bell .................................................. 241

The Datafication of Health
Minna Ruckenstein and Natasha Dow Schüll .................................. 261

China–Africa Encounters: Historical Legacies and Contemporary Realities
Helen F. Siu and Mike McGovern .................................. 337

Epidemics (Especially Ebola)
Sbaron Abramovitz .................................................. 421

Theme: Human–Animal Interaction

Humans and Animals in Northern Regions
David G. Anderson .................................................. 133
The Archaeological Study of Sacrifice
   Glenn M. Schwartz ................................................................. 223

Primate Positional Behavior Development and Evolution
   Michelle Bezanson ................................................................. 279

Archaeology and Human–Animal Relations: Thinking Through
Anthropocentrism
   Brian Boyd ................................................................................ 299

The Monkeying of the Americas: Primate Biogeography in the
Neotropics
   Jessica Lynch Alfaro ................................................................. 317

Human–Animal Communication
   Don Kulick ................................................................................. 357

Indexes

Cumulative Index of Contributing Authors, Volumes 37–46 .................. 447
Cumulative Index of Article Titles, Volumes 37–46 ................................. 451

Errata

An online log of corrections to Annual Review of Anthropology articles may be found at
http://www.annualreviews.org/errata/anthro