Book Section:

Dalmer, N., Newman-Griffis, D. orcid.org/0000-0002-0473-4226, Ibrahimi, M. et al. (4 more authors) (2024) Configuring Data Subjects. In: Jarke, J. and Bates, J., (eds.) Dialogues in Data Power: Shifting Response-abilities in a Datafied World. Bristol University Press , pp. 10-30. ISBN 9781529238303

https://doi.org/10.51952/9781529238327.ch001

1

Configuring Data Subjects

Nicole Dalmer, Denis Newman-Griffis, Mergime Ibrahimi, Xiufeng Jia, Doris Allhutter,[1] Katrin Amelang,[2] and Juliane Jarke[3]

Introduction

The collection and use of data about individual people has shifted dramatically over the last century, from a specialized practice to a near-universal norm. Once the domain of census-takers and actuaries, data about ourselves and our behaviours are now expected in nearly all situations, and we often enthusiastically collect such data about ourselves (Neff and Nafus, 2016). As data collection has changed, so too has data use: larger and richer data sets have co-evolved with techniques to analyse them and data analysis has become a part of everyday life, from the trains we ride, to the coffee we buy, to the ways government policies are designed and evaluated (LaValle et al, 2010).

But who decides what 'counts' as data, or how they are collected? And what impact do these choices have on us and the world we live in? When data are everywhere, it is easy to lose sight of the fact that they come from somewhere, and that the data we have are not the only data that could be. The ubiquity of data collection and analysis magnifies the impact of how data are conceptualized exponentially; at the same time, the increased sophistication and commercialization of data technologies have made these questions of conceptualization and design increasingly opaque. As data become ever richer, more detailed, and more incomprehensible (for example, Martin's (2019) investigation of her multi-gigabyte personal data profile from Google), there is an increasing perception that data are not just objective but perhaps even complete – that is, that given enough sensors, apps, and hard disks, we can represent an individual in toto (see Braun's (2021) discussion of this perception in the emerging science of digital twins).

This chapter challenges this perception, and brings to light the often overlooked questions of how quantification of people in data occurs and the societal and individual implications of the configuration of people as data subjects. To illustrate the multiplicity of datafication, we intentionally avoid a single definition of data subjects, allowing each contributing author to reflect individually on their interpretation of the term and its implications. Using specific examples of datafication to inform our analysis, we examine the process of datafication writ large, and what it looks like, to bring critical methods to bear on datafication and its impacts. Each section in this chapter engages with key questions in the interdisciplinary nature of how individuals are configured as data subjects in the datafication process and the methodological tools that can be brought to bear on analysing datafication, as well as important gaps in these tools that contemporary data processes highlight.

In addition to the configuration of individuals as data subjects, we put forth initial reflections on the ways in which datafication is used both to construct and to represent population groups and collective identities. While the politics of group formation and representation in data present distinct questions beyond the scope of this chapter in terms of data as a site for political enquiry and action (Beraldo and Milan, 2019), we highlight the importance of distinguishing between the datafication of populations and the datafication of individuals within those populations. Similarly, we draw out the distinctions between representing an individual's personhood and their actions, networks, or other outward spurs. There are invaluable questions shared between these scenarios – who is collecting what data, what power relationships are involved, and so on – but there are also fundamental differences between using data to represent and/or learn about the world, which population data generally aim to convey (Mooney et al, 2015), and using data to represent and/or learn about an individual (our focus in this chapter). These processes are highly interrelated, but require different methodologies as our ways of analysing data configurations continue to develop.

Examining the configuration of data subjects is both a timely and an often overlooked need. In addition to debates about digital twinning and ever-greater personalization of healthcare (Armstrong, 2017) and marketing and sales (McFall et al, 2020), among others, the ways in which people are, or fail to be, represented in data directly inform contemporary debates about responsible and ethical use of artificial intelligence (AI) (Sambasivan et al, 2021; Werder et al, 2022). Each of the following subsections further draws out the cross-disciplinary interweaving of dialogues around datafication, in disability data science, the quantified self, population ageing, and mobility. As data and datafication are increasingly implicated in all areas of the 'data-driven' society, these provide instructive examples for investigating other connections and questions in new and emerging areas.

In our writing of this chapter, we seek to equip readers with the tools to begin asking and answering questions about who becomes datafied, how this process is realized, and how datafication interacts with broader, historical debates about personhood and the self.

Design decisions and the history of data – Denis Newman-Griffis

Denis identifies as a White, non-binary, neurodivergent academic. They draw on their work on methodologies in natural language processing and AI, data science, and critical disability studies.

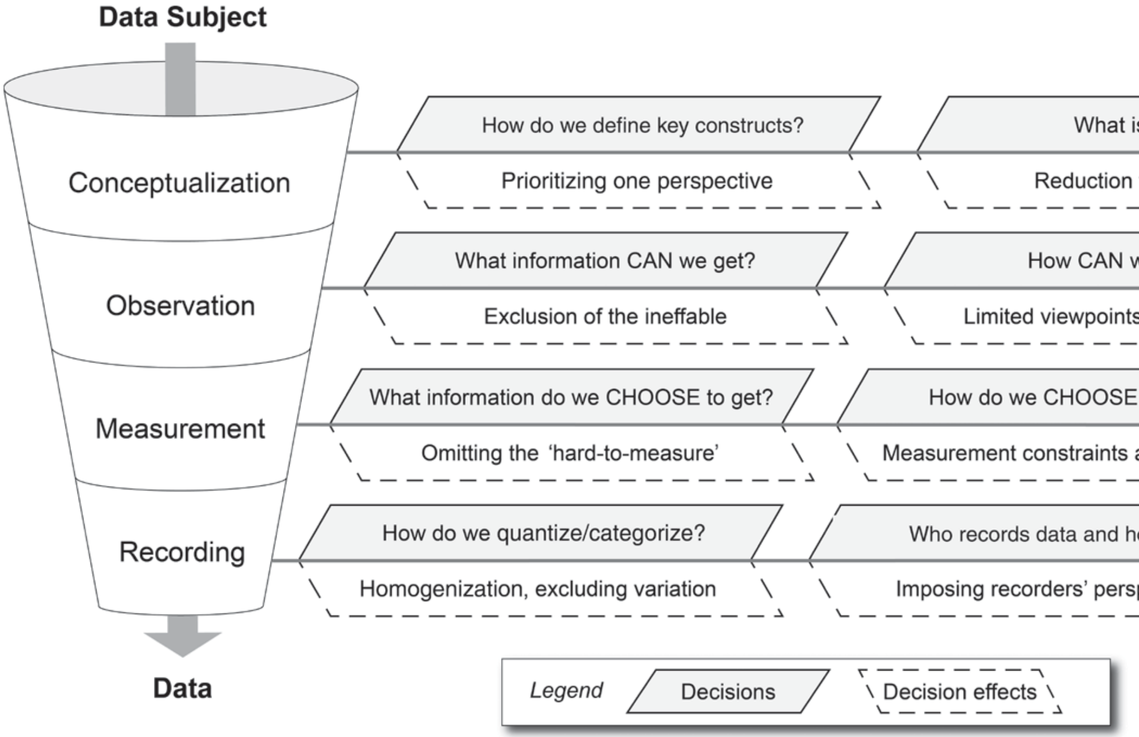
One of the first questions an informed data consumer should always ask when encountering a new data set is: Where did this come from? The companion question, which is often left unnoticed and unasked in the shadow of the enticing realms of data use, is: How did this come to be this way?

Data have both history and purpose. Measurements are taken by certain people for certain reasons; answers are elicited to specific questions from specific respondents. The purposes for which data are collected may be malicious, beneficial, or banal; sample populations may be representative or biased; data collectors may recognize and be informed by the situated nature of their work or perform it merely by rote. Each of these characteristics informs the history of data and affects its representational power and impact on the world.

More than this, however, data have conceptual history. The process of narrowing down the infinite space of information we could represent about a person to the finite subset we choose is a process of decisions, taken consciously or unconsciously over and over again each time we work with a person as data. We can imagine a data funnel representing the sequence of decisions to turn a desire for information into data that approximates that information (Figure 1.1). The act of datafication involves reframing an infinitely complex human being as a data subject, to be represented via a set of finite data. A person's lived complexity is thus approximated via a clumsy phantasm of data, and it is the decisions represented in the data funnel that configure the shape of this data ghost. Vitally, these decisions also provide anchor points for deconstructing this configuration – and for contesting it.

Consider the case of a person applying for government benefit support on the basis of disability. From an 'equal treatment' policy perspective (Mabbett, 2005), as well as from a decision consistency and accountability view (Gallicchio and Bye, 1981), the decision is typically made on the basis of standardized criteria using a fixed set of information sources to represent the person's unique disability experience. Already the process prefers information that can be directly elicited or measured and represented in codifiable ways – ineffable and subjective aspects of a person's perceptions or experience are disfavoured and likely not captured (Osterweis et al, 1987;

Figure 1.1: The data funnel: conceptual illustration of the datafication process



Note: Conceptual illustration of the datafication process by which an infinitely complex individual person, framed as a 'data subject', becomes represented in finite data. Example questions illustrating design and operational decisions at each stage of the process are shown along with some of their effects on what the data represent.

Newman-Griffis et al, 2022). Identifying what specific information to elicit, then, requires deciding: How do we choose to define disability, and what is relevant to that definition? A political and interpersonal definition necessitates different types of information from the context-focused social model, and yet different information from still-frequent medicalized views (Newman-Griffis et al, 2023).

These decisions continue down the data funnel. While it would simplify analysis, there are no single attributions of agency or motive to be made in this process: a state may require collecting information on medical conditions for statistical as well as decision-making purposes (Harrison et al, 2021); benefits agents may seek or avoid additional information within the narrow bounds afforded to them, to help an applicant or make a decision faster (Barth et al, 2017); and applicants may choose to emphasize, omit, or reframe certain information to better align with policy requirements (Halpern and Hausman, 1986). Nonetheless, power dynamics in the data collection process are clearly weighted in favour of those higher up the decision chain: a benefits applicant can do very little to answer questions that are not asked, and their answers may be omitted from recorded data as 'irrelevant' regardless.

These conceptualizations of information and decisions about its representation directly inform the design of information infrastructure, as Bowker and Star (2000) (among others) have shown. This infrastructure rapidly creates significant inertia and resistance to change: modifying data structures and contents directly impacts many of the purposes data are put to (which data actors become heavily invested in), skewing data representations heavily towards what has come before. How these implementations – and their reification in infrastructure – affect people's perceptions of one another and themselves is an intriguing and urgent question among a narrative of universal datafication. Recent work on activity tracking apps has shown that datafication of activity can come to take priority over personal experience – what you feel is what the data show, rather than vice versa (Littlejohns et al, 2019). As we further develop methods for analysing data configuration, examining the interactions between data, infrastructure, and (inter)personal perception will be rich soil to till.

As other sections of this chapter illustrate, decisions in the data funnel must be viewed in terms of the broader contexts in which those decisions are taken, and the underlying structures and perspectives that populate the players and the scene of the data collection stage. But it is tracing these decisions and reconstructing the erased historicity of data that enable us to understand how data contexts are materialized in data configurations, and to build new strategies for contesting and reimagining the configurations of our data ghosts.

Power dynamics in self-tracking data practices in the context of everyday life – Xiufeng Jia

Xiufeng identifies as a Chinese interdisciplinary researcher. She draws upon her background in AI, algorithms, and data in society, digital self-tracking, and digital health. Her current work focuses on responsible AI governance, design, and the everyday use of AI.

Data are associated with power dynamics (boyd and Crawford, 2012; Crawford et al, 2014; Van Dijk, 2014). These dynamics can involve political power, such as when data-driven technologies are used for governmental surveillance, or commercial power where data generate profits for corporations. Nonetheless, the power of data-generating individuals should not be ignored. Self-trackers, for example, create data but frequently gain relatively minimal benefits from their value (Sharon and Zandbergen, 2017). Their unique subjective experiences and narratives are crucial for understanding the value and meaning of data from their perspectives.

Ordinary individuals are configured as data subjects in self-tracking data practices, where their bodies and everyday lives are quantified into various types of data. These digital self-tracking technologies empower individuals to voluntarily collect, collate, analyse, and store their personal data. Through the quantified data, individuals strive to enhance self-knowledge and self-awareness regarding patterns in their physical exercise, food consumption, calories burned, menstruation cycles and fertility, mood, sleep, work productivity, daily expenses and other aspects of their bodies and lives in everyday life. They often rely on their personal data to understand and learn about their health conditions and bodily changes, managing their well-being either to maintain their current lifestyle and bodily situations or to improve upon them. In this sense, personal self-tracking data play a crucial role, giving people insights about how their bodies feel, what activities they engage with in their daily lives, and what decisions they need to make in their ongoing interactions with the data.

However, the intertwining of data and individuals in real-world scenarios is complex. Not all self-trackers strictly follow what their data suggest about themselves. Rather, individuals have capabilities to critically reflect on and analyse both themselves and their data in relation to their personal lived experiences and bodies. I argue that the agency of both individual and data is situated within the power dynamics of self-tracking data practices. These dynamics, often fluid and subtly present, are deeply embedded in real-life contexts. As Mascheroni (2020) suggests, understanding the everyday life contexts where data are embedded allows us to explore the power dynamics of datafication and its societal implications. Therefore, individualization and contextualization can be valuable for gaining a deeper understanding of power distribution and the agency of both individuals and data.

In particular, analysing individuals' feelings towards data and themselves presents a compelling approach to account for power relationships between data, bodies, self, and individual decision-making processes. Here, the term 'feelings' refers to emotions (for example, happiness and anxiety), bodily sensations (such as pain), and sensory experiences (for example, interactions with digital sensors). Some scholars (for example, Kennedy and Hill, 2018; Lupton et al, 2018) emphasize that people learn and behave through feelings. They suggest that feelings play an important role in understanding oneself and meanings of data to individuals. By comparing their data with their feelings, people develop embodied self-awareness and self-knowledge of their bodies.

To contextualize self-tracking data practices within the context of everyday life means to figure out individuals' feelings alongside their daily interactions with their data. Some self-tracking scholars (for example, Pantzar and Ruckenstein, 2017; Weiner et al, 2020) have stressed the importance of such contextualization for making sense of digital self-tracking data and oneself. This context becomes significant when individual self-trackers engage with their data within specific cultural and social everyday environments (Costa Figueiredo et al, 2017; Pantzar and Ruckenstein, 2017; Pink and Fors, 2017; Sharon and Zandbergen, 2017). For instance, some self-trackers feel their bodies becoming heavier, associating this with their recent dietary habits (for example, excessive intake of carbohydrates and sweets) and lifestyle changes (such as, ceasing cardio exercises). This leads them to review their relevant data such as calorie intake, weight, and physical exercise to validate their assumptions. This example demonstrates how daily habits and bodily sensations empower individuals to intuitively learn about their bodies, simultaneously indicating the powerful role of data for validating their bodily sensations, assuming they trust their data.

The interplay of the human body, data, and feelings creates a complex configuration of agency. This process shapes the construction of the self, demonstrating how individuals communicate with data and mutually engage with each other within these configurations. The ways in which self-trackers make sense of meanings of data, and use them, can be better understood within the context of their everyday lives, where they are living with their data on a regular basis.

Contextualizing the construction of older adults as data subjects – Nicole Dalmer

Nicole identifies as a queer, White settler. She draws upon her background in both Library and Information Science and critical gerontology, focusing on the role of connections (both to people and to technologies) in later life.

Digital technologies (and their associated data) are central to the shaping and (re)imagining of ageing futures. Older adults' bodies, routines, practices and behaviours are increasingly scrutinized, measured, and tracked as a means to foster healthy and active lifestyles, thus ensuring more cost-effective management of care in later life (Dalmer et al, 2022).

In exploring the multiple mechanisms that catalyse the creation of data subjects, it is imperative to identify and include those broader trends that are impacting specific population groups. The datafication of later life, for example, can be linked to four broader trends that are both contextualizing and shaping the experiences of ageing both in and with data.

The first is the imperative to age successfully. Successful ageing has gained prominence in and continues to inform theoretical paradigms, health measurements, retirement lifestyles, policy agendas, and anti-ageing ideals (Katz, 2013; Katz and Calasanti, 2015), with ageing successfully determined by: the avoidance of disease and disability, the maintenance of cognitive and physical function, and social engagement (Rowe and Kahn, 1997). Technologies (including wearables) are seen as tools to support older adults' capacity to age successfully. Digital tools (such as step counters and brain games), their data, and their 'smart' capacities can be used by older adults to self-track their performance and progress (via data points) to become more self-knowledgeable about and responsible for their well-being (Marshall, 2018). Of note, critics of this concept illuminate how successful ageing emphasizes individual responsibility (and choice) for one's lifestyle and successes (or failures) therein, without acknowledging the dynamics of power, opportunities, and inequalities.

The second broader trend is ageing in place. Ageing in place is a significant social, economic, and political goal, and refers to an individual ageing in their own home for as long as possible. Not only does ageing in place align with a majority of older adults' preferences, but it is also promoted as saving the public purse strings as it avoids costly relocation to institutional facilities. To support older adults' ability to age in place, technologies placed around the home are seen as tools to support older adults' independence in their homes and communities while simultaneously reducing (optimistically) family members' and other care partners' care work (Berridge, 2016). Consideration is less often given to the changes such digital devices (and the data they collect and transmit) bring to the feelings of privacy, security, and intimacy that are typically associated with one's home (Berridge and Wetle, 2020). The data outputted from such devices can also impact relationships, with adult children having unfettered access to and knowledge about their parents' habits and comings and goings. And so, while older adults prefer to age in their own homes, the tensions that digital devices in the home (that surveil, monitor, and report to others) introduce cannot be ignored.

The third trend, exacerbated by COVID-19, is the growing concern regarding social isolation among older adults. Older adults' experiences of social isolation and loneliness are an increasingly important topic of conversation in many countries, in part due to the many negative impacts of social isolation on older adults' physical and mental health and well-being: increased risk of premature death, depression, falls, cardiovascular disease, and dementia (Courtin and Knapp, 2017; Sepúlveda-Loyola et al, 2020). As evidence of an interventionist logic (Peine and Neven, 2019), technologies and associated transmission of data are proposed as interventions ('solutions') to solve the 'problems' of lonely or isolated older adults. Older adults are increasingly encouraged to share their personal data (via social media sites, email, photo sharing, and video chat tools) in order to improve their feelings of social connection (Petersen et al, 2023), thus also ensuring their capacity to age successfully.

Taken together, the imperative to age successfully and age in place, combined with the fears of social isolation in later life are changing the ways that bodies, homes, communities, and other spaces of care are understood; reframed as domains to be monitored, tracked, and managed via data to ensure that older adults' well-being and health are supported. In other words, we can see how these trends mould older adults as subjects of systems of governance.

Curiously, while the objective to support older adults to age cost effectively and successfully in place relies on the gathering and tracking of data, older adults themselves are often excluded from data gathering practices. Accordingly, the fourth and final trend that both creates and shapes older adults as data subjects are the many ageist stereotypes that are held about later life. Fernández-Ardèvol and Grenier (2022) have recently proposed the concept of data ageism[4] to convey the production and reproduction of the disadvantaged status of older age as a result of decisions concerning how, when, and where to collect and deliver data. Older adults, in data collection practices, are often treated as residual categories, which not only mutes the heterogeneity inherent in older adult populations, but can reinforce the invisibility of some older adults (Fernández-Ardèvol and Grenier, 2022). Given that data shape 'how and what we see' (Sendyka, 2013: 104), such exclusions can ultimately contribute to or perpetuate ageist thinking and practices.

Who is a (data) subject for autonomous vehicles? – Mergime Ibrahimi

Mergime identifies as a woman whose background is in critical data science, focusing on mobility, data diversity and perceptions towards automated systems.

Techno-companies and institutions that use technology often portray innovative data-driven technologies in a positive light, emphasizing their objectivity and efficiency in managing processes in different domains. Our understandings and imaginaries about autonomous vehicles (AVs), for example, are derived mostly from car companies' advertisements and thus create imaginary assumptions in the public that AVs will shape the future of mobility (Martin, 2021), and will solve the world's mobility problems. However, the ways in which people are represented (or fail to be represented) in data create contemporary debates about responsible and ethical use of AI. This perspective invites us to critically examine who gets to be defined as a relevant data subject for AVs.

Considering that AVs are envisaged to operate autonomously in complex and diverse societal settings, the configuration of data subjects becomes a critical question: Who is defined as a data subject – the passenger(s) inside the vehicle or other individuals in the surrounding environment? And how are the other individuals categorized and classified? On one hand, the passengers sitting inside the vehicle are the direct users of the technology who are subject to the decisions made by the AV. The configuration of these data subjects revolves around making sure of their safety, comfort, and convenience during transportation. On the other hand, there is the surrounding environment – other individuals around AVs such as pedestrians, cyclists, delivery robots, and other vehicles who need to both share the space with AVs and feel safe around them. It is AV designers who determine and define what encompasses the 'relevant' surrounding environment of AVs based on the input data they choose to incorporate into the algorithms. The AV algorithms are designed to process vast amounts of data rooted in a historical trajectory shaped by societal values, norms, and past technological advancements. Historical inequalities and biases can be inadvertently perpetuated in the data used to train these algorithms (Liu, 2017; Lim and Taeihagh, 2019), influencing how AVs perceive and respond to their surroundings. The lack of representation and agency for certain groups in historical data can have cascading effects on the configuration of data subjects in AV systems. Studies have revealed that AVs often fail to 'see' wheelchair riders or individuals with characteristics not present in the training data (Whittaker et al, 2019), as data diversity is often ignored. Also, research has identified potential biases within machine vision in general and the specific pedestrian detection algorithms as it tends to perform less accurately on females and individuals with darker skin tones (Buolamwini and Gebru, 2018), mirroring a long-standing history of hostile architecture and urban planning that rarely considers the needs of people with mobility impairments or marginalized communities.

One of the ethical challenges arising from identifying potential victims of car accidents is the determination of priorities. AVs must make split-second decisions when there might be an error in the system in potentially dangerous situations, such as choosing between crashing with an object or swerving to avoid it, which puts pedestrians or cyclists at risk. These decisions are based on complex algorithms trained on large datasets, but the lack of transparency in the training process and the specific data used raises questions about fairness, accountability, and the potential for bias.

Should AVs prioritize the safety of the passengers, considering the idea that users are paying for protection? Or should AVs prioritize the safety of pedestrians or other road users, upholding the principle of minimizing overall harm? Or should AVs save only the people whose data dominate the training datasets and eventually push certain other users off the roads? This conundrum extends beyond mere algorithmic calculations; it delves into the realm of moral and ethical reasoning (Robinson et al, 2022) and how that is perpetuated in AV algorithms.

Critically, this section does not propose that the overall goal of future work should be to simply lower the barriers to collecting and filtering the data. Rather, this section highlights a swath of normative questions about who becomes a data subject for AVs and which context data are considered essential in configuring these subjects. It has been argued that even if we use all available data, we still can get biased results because society is biased (Caliskan et al, 2017). This is because data-driven technologies are trained with data that tend to carry on the social dynamics and iterate 'the patterns of marginalisation, inequality, and discrimination' that exist in our societies, and are thus represented in the data (Leslie, 2019: 4). However, biases in data are not only generated by the discriminations that exist in society but rather by the contexts they are drawn upon, and the assumptions made during algorithm development. The filtering and selection of data plays a crucial role in shaping the outcomes and configurations of data subjects within AVs and other data-driven technologies.

In short, only those who are represented in the training data sets and have been historically considered 'relevant' are more likely to become data subjects that AV algorithms prioritize and consider in decision-making processes. This perpetuates a cycle of underrepresentation and exclusion, and highlights the need for a critical approach that considers the concerns and expectations of a broader range of data subjects, who often have very little agency in processes of technological governance.

Concluding remarks

The history, purposes, and contexts of data formation, collection, and use play fundamental roles in shaping the power and impact of data and data technologies. As we illustrate in this chapter, power dynamics are inextricably embedded in everyday data practices and in the configuration of individuals as data subjects.

The (re)configuring of people as data subjects within datafied systems and our broader datafied society is a complex process reflecting layered, multi-party histories of data construction and capture, design decisions and assertions of power, and technical implementation together with individual perception. Teasing out the nuances of power dynamics, notions and exercise of agency, and historical and contextual factors that underpin datafication requires multidimensional, transdisciplinary approaches. This chapter illustrates how such boundary-spanning approaches are vital to elucidate questions of how quantification of people in data occurs, and what the implications are for individuals and society.

Context plays a vital role in defining and shaping data subjects. The power dynamics within datafied systems are entwined with diverse and heterogeneous sociocultural environments, and these environments and the people within them shape every stage of data definition, collection, materialization, and use. Measuring people and perception based solely on 'universal' standards inevitably overlooks important contextual nuances, limiting our understanding of the impact of datafication on different communities and disparities in their input to datafication processes. By examining the feelings and experiences of individuals in their specific contexts, we can gain a deeper understanding of the power relationships between data, bodies, the self, and individual decision-making.

In each section of this chapter, we have invited the reader to question who becomes datafied, who enacts datafication, how this process is realized, and how datafication interacts with broader historical debates about personhood and the self. Taking up the findings from this chapter in conversation with the concept of the data funnel as illustrated in Figure 1.1: Denis' introductory section on the history of data begins to illuminate how the ways in which we define data and determine relevance come to bear on the *conceptualization* of data subjects; both Xiufeng's exploration of self-tracking practices and Nicole's broader examination of the datafication processes of a specific population (older adults) reveal the important questions that must be asked and decisions that must be made in the *conceptualization* and *observation* processes of the creation of the data subject (and how these decisions are materialized in data *recording*); and finally, Mergime's section critically examines who become data subjects for AVs and how the inherent biases and exclusions of potential data subjects are perpetuated by historical representation and relevance in training data sets.

This chapter reflects the multiplicity of datafication, both in the multiple perspectives that inform the representation of a single individual in data and in the multiple methodologies needed to critically examine and question the configuration of data subjects. To effectively engage with the process and politics of datafication, and the interwoven systems that configure individual data subjects, it is crucial to consider the historical context, power dynamics, and ethical implications that shape the collection, interpretation, and representation of data. People's individual perspectives, concerns, and expectations shape the power dynamics between themselves as data subjects and the datafication solutions that aim to encapsulate them. Data experts, as the creators and gatekeepers of data systems and the algorithms to analyse them, shape the data landscape. And insights into the often-inscrutable processes of algorithmic manipulation of data and data-driven decision-making is vital to understand how algorithms materialize datafication and its impacts. Examining the interplay between data subjects and data experts, the algorithms they create and are affected by, and the contexts they emerge from, operate in, and shape, are at the heart of understanding how people are configured as data subjects and imagining alternative (re)configurations.

DISCUSSANT RESPONSES

Individual and socio-technical practices of configuring data subjects – Katrin Amelang

Katrin responds to this nicely curated compilation deconstructing data subjects from the perspective of a science and technology studies inspired cultural anthropologist from Germany researching human–technology relations. In particular, she brings to the dialogue her recent work on the datafication of health and the body.

In critical data studies (CDS), one of the perhaps well-worn but also well-won commonplaces is that data are never raw or objective, merely depicting a neutral image of the world, but a specific, both partial and powerful way of capturing, knowing, and narrating it. Hence, CDS research emphasizes that digital data (sets) – like other results of knowledge production – have to be situated in particular times and places as well as in specific (data) practices and material-semiotic arrangements (for example, Loukissas, 2019; Kitchin, 2021). Against this backdrop and based on their respective research fields, the four authors focus on how individuals are made knowable in and through data, and thus on datafication processes, in which people are translated into digital data and are shaped as subjects of these data. While my research on mobile apps for menstrual self-observation provides another example supporting the authors' analysis of configuration processes of data subjects, I want to highlight and follow up on the socio-technical co-production of individuals/data subjects and data technologies.

The authors pay particular attention to the manifold conceptual choices made in the generation and collection of data and how these affect the configuration of data subjects. In this respect, Denis starts off with figuring datafication with the data funnel as decision and reduction processes, in which an individual framed as a specific data subject is narrowed down for the purpose of (specific) representation. All of them show such decisions are neither neutral nor random and anything but innocuous in their implications. For example, Nicole elaborates how sociocultural ideas, norms, stereotypes, and imaginaries about later life not only define older age as a specific problem and shape data collection and technologies designed for it, but also how older adults are grouped and homogenized on this basis, regardless of their different circumstances of life. Further, Mergime illustrates the conceptual power and historical legacies in the inclusion and exclusion of certain groups of people that, in her case of AVs, prioritize who becomes or does not become a data subject, and determine what becomes training data and thus implemented in technical innovations.

Adding the case of period-tracking, it is interesting who is being addressed as a menstruating person, potential app user, and data subject as well as how they are being imagined to be. The tracking categories these apps offer structure menstrual-self-monitoring practices and define what can (and cannot) be counted, tracked, and become data. The way apps engage users through user interface design and push-up notifications reveals their socio-technical scripts, which say as much about computational attempts at calculating menstrual cycles as they do about sociocultural ideas about menstruation and menstruating people or the politics of gender, sexuality, and reproduction, not to mention the politics and value of (menstrual) data. While the examples of all authors make clear that the conceptual history of data is inscribed, reflected, and materialized in the design of data systems, I would like to emphasize this point so as not to lose sight of the at once semiotic/social and material/technical constellations and configurations in datafication processes. Or, in Suchman's (2007) sense, data, data technologies, and data subjects are in a relationship of mutual constitution (and thus change).

Individuals are assembled and configured as data subjects not only by design through power-laden conceptualizations and realizations of data generation but also by encountering and using these data applications. Conversely, data technologies evolve with their uses and users. Not all data subjects are active users of data systems and not all actors putting data to use are the subject of these data (see Nicole's example of older adults). In this regard, the authors hint at the power of cooperative actors in datafication processes or indicate that options for decision and action are often not clear-cut and easily assignable. Xiufeng, in particular, deals with the power dynamics that unfold between data and individuals in the realm of self-tracking practices. In line with Sharon and Zandbergen (2017), Pantzar and Ruckenstein (2017), and Weiner et al (2020), she argues for embedding data practices in everyday life, considering meanings and values of data that not only reproduce but also challenge hegemonic notions of objective data truths, as well as paying more attention to the reflexive interplay of body data and feelings.

Again, I would emphasize technology (in this case, the app) in this far from unambiguous and frictionless interplay where people reconcile their datafied and embodied selves. The sensemaking of individual body experiences via app unfolds in a field of tension between normative pre-configurations (default settings) and moments of self-empowerment. In my conversations with users of period-tracking apps, the contingencies and contradictions of apps as instruments supporting self-knowledge become clear, but also how they deal with the ambivalent effects, gains, and uncertainties of data technologies in wayward, pragmatic, and reflexive ways (Amelang, 2022). Agential possibilities of data subjects may be more obvious in the case of individual, data-generating practices of self-tracking. Yet even in cases where individuals do not fit into, or are inadequately represented in, data systems (see Denis's example of applying for government benefits for the disabled), practices of failing, fitting into, or attempting to outsmart systems can be understood as part of the (re)configuration of individuals as data subjects. Emphasizing users does not mean forgetting other actors and data uses or underestimating the power of corporate and state actors, but does remind of the frictions and ambivalences in datafication processes (Hepp et al, 2022).

All this being said, I appreciate the deconstructing and critical reflexive approach of the authors. By exploring the question of who becomes a data subject and how, they introduce readers to various moments of both the construction of data and the configuration of data subjects. They thus also provide a starting point for further empirical research and for thinking how configurations of data subjects could be otherwise. This should include the role of technology in configuring the relation of data (subjects) and 'fleshy selves' as well as the moving beyond notions of accurate representation. Last but not least (and bridging to Doris), one should not forget that datafication always entails the promise of calculation, which translates between and relates individuals with populations and configures data subjects accordingly.

The inseparability of the individual from the collective – Doris Allhutter

Doris believes in the impact of collective subjectivity on research, but never came up with something that made sense to her when asked to identify. She writes this response with an interest in the entanglement of individual and collective subjectivity.

The authors of this chapter trace how datafication and data practices affect individuals and call to examine the configuration of data subjects through the lens of power dynamics, agency, and historical context. While reading the contributions, I wondered why it seems important to distinguish between the datafication of populations and individuals when historically grown power relations necessarily pervade the macro-, meso-, and micro-levels of society. My response attempts to trace and compare the power dynamics that the authors tease out in their respective cases.

Asking who gets to be defined as a relevant data subject for AVs, Mergime Ibrahimi points out the historical neglect of the needs of people with impairments and marginalized communities in urban planning and mobility (see Wilson et al, 2019). Privileges of safety and convenience may first go to solvent buyers of AVs. Technological governance, Mergime suggests, needs to be more inclusive and extend unequally distributed agencies to underrepresented communities. However, power would still side with the industry's individualistic vision of the future of mobility as private transport, as Mergime hints to in her introduction of the case. A moral dilemma perspective that ignores public value and argues case-by-case tolerates the ableist, racist, and classist power imbalances at play, I suggest.

Nicole Dalmer describes how older adults are constructed as subjects of governance systems to ensure more cost-effective management of care. She remarks that power dynamics influence people's capabilities of self-responsibility in 'successful ageing in place' and implicitly indicates intersectional identity categories such as gender, class, race, and ability. A critical ambivalence that points to restrictions on older adults' agency shows in the contradictory demands imposed on them. On the one hand, social isolation is seen as a matter of health concern that ought to be self-managed by maintaining social connections via social media. On the other hand, prioritizing self-responsibility and autonomy over care promotes that successful, data-supported ageing in place is an isolated activity. This individualization seems justified by the idea of a 'collective advantage-through-technology' (see Öchsner, 2021).

Xiufeng Jia describes self-trackers as data-generating individuals who aim to enhance their self-knowledge about data patterns derived from their bodies and daily practices. She defines power and agency as distributed between individuals and data and states that analysing individuals' feelings towards data and themselves accounts for power relations between data and self. Thinking with queer-feminist theories of affect and emotion, political theorist Brigitte Bargetz (2019) shows how affects and emotions point to the way in which people are imbricated in power relations. Affects and emotions do not indicate a subjective state, and they do not stand outside the social. Inequality produces specifically feeling subjects. Since data practices accommodate some people, bodies, and everyday practices more than others, they affect subjects differentially (Allhutter, 2021). We can start by asking who is invited to gain self-knowledge about their bodies and their everyday. And how do ideologies of human difference frame what kind of self-tracking practices address whom?

Finally, Denis Newman-Griffis focuses on the conceptual history of data. Data collected from a person applying for disability benefits mirrors a medicalized view of disability. It hardly considers information informed by the context-focused social model representing the person's disability experience. Down the data funnel, agency is distributed between the state, the benefits agents, and the applicants to ask and give information that influences the decision outcome, to dismiss some as irrelevant or apply discretion. Denis highlights the data's representational power, which affects people's perceptions of one another and themselves, and thus emphasizes the agentive historicity of data configurations, which are black-boxed by the erasure of the contexts of their emergence. Power also shows in the way representational bias is intimately linked with the material conditions of possibility for inclusion and exclusion.

The concluding remarks of the chapter focus the reader's attention on 'the power relationships between data, bodies, the self, and individual decision-making' by 'examining the feelings and experiences of individuals in their specific contexts'. This is a compelling approach. My reading tries to emphasize the richness of the authors' mobilization of power-critical perspectives. While centring on the configuration of individual data subjects, they offer a variety of entrance points to unpack power dynamics that operate through multiple layers of society. The datafication of populations and the datafication of individuals are powerful because they are inseparable.

Notes

[1] Discussant

[2] Discussant

[3] Facilitator

[4] Related, in part, to digital ageism, or age biases present in technologies, such as AI, where older adults may be excluded from technology development and design considerations as a result of prejudices or stereotypes (Manor and Herscovici, 2021; Chu et al, 2022).

References

Allhutter, D. (2021) 'Memory Traces in Society–Technology Relations: How to produce cracks in infrastructural power'. In: R. Hamm (ed) *Reader Collective Memory-Work*. Sligo: BeltraBooks, pp 426–452.

Amelang, K. (2022) '(Not) Safe to Use: Insecurities in everyday data practices with period-tracking apps'. In: A. Hepp, J. Jarke, and L. Kramp (eds) *New Perspectives in Critical Data Studies: The Ambivalences of Data Power.* Basingstoke & London: Palgrave Macmillan, pp 297–321.

Armstrong, S. (2017) 'Data, data everywhere: the challenges of personalised medicine', *BMJ*, 359.

Bargetz, B. (2019) 'A Political Grammar of Feelings: Thinking the political through sensitivity and sentimentality'. In: T. Bedorf and S. Herrmann (eds) *Political Phenomenology*. New York: Routledge, pp 52–70.

Barth, J., de Boer, W. E., Busse, J. W., Hoving, J. L., Kedzia, S., Couban, R., et al (2017) 'Inter-rater agreement in evaluation of disability: systematic review of reproducibility studies', *BMJ*, 356.

Beraldo, D. and Milan, S. (2019) 'From data politics to the contentious politics of data', *Big Data & Society*, 6(2): 2053951719885967.

Berridge, C. (2016) 'Breathing room in monitored space: the impact of passive monitoring technology on privacy in independent living', *The Gerontologist*, 56(5): 807–816.

Berridge, C. and Wetle, T. F. (2020) 'Why older adults and their children disagree about in-home surveillance technology, sensors, and tracking', *The Gerontologist*, 60(5): 926–934.

Bowker, G. C. and Star, S. L. (2000) *Sorting Things Out: Classification and its Consequences*. Cambridge, MA: MIT Press.

boyd, D. and Crawford, K. (2012) ‘Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon’, *Information, Communication & Society*, 15(5): 662–679.

Braun, M. (2021) ‘Represent me: please! Towards an ethics of digital twins in medicine’, *Journal of Medical Ethics*, 47(6): 394–400.

Buolamwini, J. and Gebru, T. (2018) ‘Gender shades: intersectional accuracy disparities in commercial gender classification’. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR (Proceedings of Machine Learning Research), 81: 77–91.

Caliskan, A., Bryson, J. J., and Narayanan, A. (2017) ‘Semantics derived automatically from language corpora contain human-like biases’, *Science*, 356(6334): 183–186. doi: 10.1126/science.aal4230. PMID: 28408601.

Chu, C. H., Nyrup, R., Leslie, K., Shi, J., Bianchi, A., Lyn, A., et al (2022) ‘Digital ageism: challenges and opportunities in artificial intelligence for older adults’, *The Gerontologist*, 62(7): 947–955.

Costa Figueiredo, M., Caldeira, C., Reynolds, T. L., Victory, S., Zheng, K., and Chen, Y. (2017) ‘Self-tracking for fertility care: collaborative support for a highly personalized problem’. *Proceedings of the ACM on Human–Computer Interaction*, 1(CSCW): 1–21.

Courtin, E. and Knapp, M. (2017) ‘Social isolation, loneliness and health in old age: a scoping review’, *Health & Social Care in the Community*, 25(3): 799–812.

Crawford, K., Gray, M. L., and Miltner, K. (2014) ‘Critiquing big data: politics, ethics, epistemology’. Special section introduction. *International Journal of Communication*, 8: 1663–1672.

Dalmer, N., Ellison, K., Katz, S., and Marshall, B. (2022) ‘Ageing, embodiment and datafication: dynamics of power in digital health and care technologies’, *International Journal of Ageing and Later Life*, 15(2): 77–101.

Fernández-Ardèvol, M. and Grenier, L. (2022) ‘Exploring data ageism: what good data can(’t) tell us about the digital practices of older people?’, *New Media & Society*, 14614448221127261.

Gallicchio, S., and Bye, B. (1981) Consistency of initial disability decisions among & within states. No. 13. US Department of Health and Human Services, Social Security Administration, Office of Policy, Office of Research and Statistics.

Halpern, J. and Hausman, J. A. (1986) ‘Choice under uncertainty: a model of applications for the social security disability insurance program’, *Journal of Public Economics*, 31(2): 131–161.

Harrison, J. E., Weber, S., Jakob, R. and Chute, C. G. (2021) ‘ICD-11: An International Classification of Diseases for the Twenty-First Century’, *BMC Medical Informatics and Decision Making*, 21(6): 206. doi: 10.1186/s12911-021-01534-6.

Hepp, A., Jarke, J. and Kramp L. (2022) 'New Perspectives in Critical Data Studies: The Ambivalences of Data Power – An Introduction'. In: A. Hepp, J. Jarke, and L. Kramp (eds) *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*. Basingstoke & London: Palgrave Macmillan, pp 1–23.

Katz, S. (2013) 'Active and successful aging: lifestyle as a gerontological idea', *Recherches sociologiques et anthropologiques*, 44(44–1): 33–49.

Katz, S. and Calasanti, T. (2015) 'Critical perspectives on successful aging: does it "appeal more than it illuminates"?', *The Gerontologist*, 55(1): 26–33.

Kennedy, H. and Hill, R. L. (2018) 'The feeling of numbers: emotions in everyday engagements with data and their visualization', *Sociology*, 52(4): 830–848.

Kitchin, R. (2021) *Data Lives: How Data Are Made and Shape Our World*. Bristol: Bristol University Press.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., and Kruschwitz, N. (2010) 'Big data, analytics and the path from insights to value', MIT Sloan Management Review.

Leslie, D. (2019) 'Understanding artificial intelligence ethics and safety', arXiv preprint, arXiv:1906.05684.

Lim, H. S. M. and Taeihagh, A. (2019) 'Algorithmic decision-making in AVs: understanding ethical and technical concerns for smart cities', *Sustainability*, 11(20): 5791.

Littlejohns, R., Gouthro, M. B., and Dickinson, J. (2019) 'Runners' engagement and social support practices: exploring the uses and role of online activities', *Sport in Society*, 22(12): 2243–2260. doi: 10.1080/17430437.2019.1571486.

Liu, H. Y. (2017) 'Three types of structural discrimination introduced by autonomous vehicles', *University of California Davis Law Review Online*, 51: 149–180.

Loukissas, Y. A. (2019) *All Data Are Local: Thinking Critically in a Data-driven Society*. Cambridge, MA: MIT Press.

Lupton, D., Pink, S., Heyes LaBond, C., and Sumartojo, S. (2018) 'Digital traces in context: Personal data contexts, data sense, and self-tracking cycling', *International Journal of Communication*, 12: 647–666.

Mabbett, D. (2005) 'Some are more equal than others: definitions of disability in social policy and discrimination law in Europe', *Journal of Social Policy*, 34(2): 215–233. doi: 10.1017/S0047279404008554.

Manor, S. and Herscovici, A. (2021) 'Digital ageism: a new kind of discrimination', *Human Behavior and Emerging Technologies*, 3(5): 1084–1093.

Marshall, B. L. (2018) 'Our Fitbits, Our (Ageing) Selves: Wearables, Self-Tracking and Ageing Embodiment'. In: S. Katz (ed) *Ageing in Everyday Life*. Bristol: Policy Press, pp 197–214.

Martin, N. (2019) 'How Much Does Google Really Know About You? A Lot', Forbes, 11 March. Available at: www.forbes.com/sites/nicolemartin1/2019/03/11/how-much-does-google-really-know-about-you-a-lot/?sh=73d856bb7f5d (accessed 16 July 2023).

Martin, R. (2021) 'AV futures or futures with AVs? Bridging sociotechnical imaginaries and a multi-level perspective of autonomous vehicle visualizations in praxis', *Humanities and Social Sciences Communications*, 8(1): 1–15.

Mascheroni, G. (2020) 'Datafied childhoods: contextualising datafication in everyday life', *Current Sociology*, 68(6): 798–813.

McFall, L., Meyers, G., and Hoyweghen, I. V. (2020) 'The personalisation of insurance: data, behaviour and innovation', *Big Data & Society*, 7(2): 2053951720973707.

Mooney, S. J., Westreich, D. J., and El-Sayed, A. M. (2015) 'Commentary: epidemiology in the era of big data', *Epidemiology*, 26(3): 390–394.

Neff, G. and Nafus, D. (2016) *Self-tracking*. Cambridge, MA: MIT Press.

Newman-Griffis, D. R., Hurwitz, M. B., McKernan, G. P., Houtrow, A. J., and Dicianno, B. E. (2022) 'A roadmap to reduce information inequities in disability with digital health and natural language processing', *PLOS Digital Health*, 1(11).

Newman-Griffis, D. R., Rauchberg, J. S., Alharbi, R., Hickman, L., and Hochheiser, H. (2023) 'Definition drives design: disability models and mechanisms of bias in AI technologies', *First Monday*, 28.

Öchsner, S. (2021) Is all ok? AAL as infrastructure for aging well. Doctoral thesis at the University of Vienna, Department of Science and Technology Studies.

Osterweis, M., Kleinman, A., and Mechanic, D. (1987) 'Conflicts and Contradictions in the Disability Program'. In: M. Osterweis, A. Kleinman, and D. Mechanic (eds) *Pain and Disability: Clinical, Behavioral, and Public Policy Perspectives*. Washington, DC: National Academies Press (US), pp 66–86.

Pantzar, M. and Ruckenstein, M. (2017) 'Living the metrics: self-tracking and situated objectivity', *Digital Health*, 3. doi: 10.1177/2055207617712590.

Peine, A. and Neven, L. (2019) 'From intervention to co-constitution: new directions in theorizing about aging and technology', *The Gerontologist*, 59(1): 15–21.

Petersen, B., Khalili-Mahani, N., Murphy, C., Sawchuk, K., Phillips, N., Li, K. Z. H., and Hebblethwaite, S. (2023) 'The association between information and communication technologies, loneliness and social connectedness: a scoping review', *Frontiers in Psychology*, 14. doi: 10.3389/fpsyg.2023.1063146.

Pink, S. and Fors, V. (2017) 'Being in a mediated world: self-tracking and the mind–body–environment', *Cultural Geographies*, 24(3): 375–388.

Robinson, J., Smyth, J., Woodman, R., and Donzella, V. (2022) 'Ethical considerations and moral implications of autonomous vehicles and unavoidable collisions', *Theoretical Issues in Ergonomics Science*, 23(4): 435–452.

Rowe, J. W. and Kahn, R. L. (1997) 'Successful aging', *The Gerontologist*, 37(4): 433–440.

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., and Aroyo, L. M. (2021) '"Everyone wants to do the model work, not the data work": data cascades in high-stakes AI', *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp 1–15.

Sendyka, R. (2013) 'Scopic regimes and modernity: hypotyposis'. In: D. Koczanowicz and D. Schauffler (eds) *Discussing Modernity: A Dialogue with Martin Jay*. Amsterdam & New York: Brill, pp 103–114.

Sepúlveda-Loyola, W., Rodríguez-Sánchez, I., Pérez-Rodríguez, P., Ganz, F., Torralba, R., Oliveira, D. V., and Rodríguez-Mañas, L. (2020) 'Impact of social isolation due to COVID-19 on health in older people: mental and physical effects and recommendations', *The Journal of Nutrition, Health & Aging*, 24: 938–947.

Sharon, T. and Zandbergen, D. (2017) 'From data fetishism to quantifying selves: self-tracking practices and the other values of data', *New Media & Society*, 19(11): 1695–1709.

Suchman, L. (2007) *Human-machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press.

Van Dijk, T. A. (2014) *Discourse and Knowledge: A Sociocognitive Approach*. Cambridge: Cambridge University Press.

Weiner, K., Will, C., Henwood, F., and Williams, R. (2020) 'Everyday curation? Attending to data, records and record keeping in the practices of self-monitoring', *Big Data & Society*, 7: 1–15.

Werder, K., Ramesh, B., and Zhang, R. (2022) 'Establishing data provenance for responsible artificial intelligence systems', ACM Transactions on Management Information Systems (TMIS), 13(2): 1–23.

Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., et al (2019) 'Disability, bias, and AI', *AI Now Institute*.

Wilson, B., Hoffman, J., and Morgenstern, J. (2019) 'Predictive inequity in object detection', arXiv preprint, arXiv:1902.11097.