The Backend Work of Data Subjects: Ordinary Challenges of Living with Data in India and the US

RANJIT SINGH, *Data & Society Research Institute*

Suggested Citation:

Ranjit Singh, 'The Backend Work of Data Subjects: Ordinary Challenges of Living with Data in India and the US', in Lisa Parks, Julia Velkova, and Sander de Ridder (eds.), *Media Backends: The Politics of Infrastructure, Clouds, and Artificial Intelligence* (Illinois: University of Illinois Press, 2023). pp. 229-244.

Abstract:

Just like data systems have a frontend and a backend, the work that data subjects must do to live with data systems also has a frontend and a backend demarcated by their activities in relation to the digital interface. This chapter uses digital interfaces as an organizing principle to analyze the unevenly distributed work of living with data systems. Using examples from everyday experiences of data subjects in India and the United States, it shows that the mutual shaping of data records and everyday lives of data subjects happens not only at the digital interface, but also around them in the backend and relies on conviviality of human interfaces to data systems. This backend work is the uneven cost of living in a data-driven world.

Keywords:
Data Subjects, Digital Interfaces, Human Interfaces, Backend work, Frontend Work

Introduction: Becoming a Data Subject

In this chapter, I analyze the work that people who are subject to data-driven practices, or data subjects[1], must do to secure representation in and contend with the consequences of data systems in everyday life. I will show how this work is an emergent condition of living in a data-driven world. I use interfaces *of* and *to* data systems as the organizing principle for my analysis. By interface *of* data systems, I imply digital interfaces that enable data subjects to interact with data systems. By interfaces *to* data systems, I am referring to a range of intermediaries from professionals (customer service operators, street-level bureaucrats, consultants, etc.) to individuals in support networks (friends and family, online and offline data subject communities, etc.) who often act as proxies for data subjects and interact with digital interfaces of data systems on their behalf. In short, I approach interfaces as both digital and human touchpoints of access to data systems. While working *on* digital interfaces has

[1] Critical data studies has increasingly moved away from the figure of the 'user' (Woolgar 1990; Oudshoorn and Pinch 2003; Hyysalo, Jensen, and Oudshoorn 2016) to analytically articulate the position of people who are increasingly made subject to data systems without their consent, participation, or, at times, even knowledge (Couldry and Yu 2018), especially in the context of automated scoring systems such as credit scores in the United States (Ziewitz and Singh 2021). Data subjects is used instead to highlight the dual role that people play as resources as well as targets of data systems in regulations (GDPR 2018) and by critical data studies scholars to describe people who live with data systems. Data "subjectifies through practices of production, accumulation, aggregation, circulation, valuation, and interpretation. These practices call upon subjects who are not separate from but submit to and are active in the various ways that data is made [...]. People govern their health by making themselves data subjects of health" (Ruppert, Isin, and Bigo 2017, 5).

always been constitutive of the lived experience of data subjecthood, working *with* human interfaces is also increasingly becoming a precondition for data subjects to make data systems work for them.

This focus on interfaces borrows part of its inspiration from the focus of this edited volume on media backends, which I conceptually approach as a site of struggle over understanding the workings of data systems and their profound consequences for data subjects. This struggle is not only about contending with opacity of data systems (Burrell 2015), but also about aligning the data produced and maintained by data systems with an individual data subject's way of life. In this sense, the interface binds the frontend and backend of data systems into an emergent relationship. On one hand, the frontend is a data system's data subject-facing side. It is often associated with what is visible, accountable, presentable, and usable about a data system. A typical example of frontend is the digital interface of any data system. On the other hand, the backend is where a data system is made workable, away from the gaze of data subjects. It sinks into the background, becomes invisible, and only returns to conscious reflection when data systems do not work as expected or break down. Typical examples where backend becomes visible is calling a customer service operator (engaging with human interfaces) for a digital service to override/renegotiate predefined terms and conditions of navigating its digital interface. The frontline human interfaces are the infrastructure (Simone 2004) of the digital interface.

The other part of the inspiration comes from the central role that the digital interface plays in how data subjects and data systems mutually shape each other. While the function of the digital interface in demarcating frontend and backend of data systems is well understood, I want to draw attention to how the digital interface is a useful analytic resource to draw out demarcations between the different kinds of work that data subjects must do to make data systems work for them. My interest in these kinds of work began quite early, when I met Yogita[2], one of my first field respondents, during fieldwork[3] for my dissertation research on Aadhaar[4] (translation: foundation), India's biometrics-based national identification infrastructure, in the summer of 2015. In narrating her troubles in obtaining a marriage registration certificate without Aadhaar, she concluded:

> Now when you go to these [government] offices, people have found a new excuse for why they cannot do your work. *Computers are the new 'babus'[5]!* They will tell you things like: 'Madam, we want to register your marriage, but this computer won't let us'! (Yogita, personal communication, 3 August 2015, emphasis added)

[2] All respondents have been anonymized and their affiliations masked to protect their privacy.

[3] The fieldwork was conducted in three rounds: between June 2015 and Jan. 2016; between July 2016 and Jan. 2017; and between Jan. 2018 and Mar. 2018. The fieldwork was multi-sited (Marcus 1995) including locations such as startup workspaces in Bengaluru; and offices of the Unique Identification Authority of India, other Aadhaar-related service outlets, the Supreme Court of India, and activist organizations in Delhi. I conducted more than 100 semi-structured qualitative interviews in English and Hindi with Supreme Court lawyers, activists, Aadhaar technology designers, NGO representatives involved in helping residents use Aadhaar, and finally, residents navigating the different key processes of Aadhaar.

[4] Taking the length of this chapter into account, I only provide information about the design and implementation of Aadhaar, which is relevant to understand my field stories. For a more detailed account of the project, see (Aiyar 2017; Nilekani and Shah 2015) and its critiques, see (Khera 2018).

[5] Hindi for 'a street-level bureaucrat' or 'a government servant'. While its semiotic analysis is outside the scope of this chapter, the term 'babu' exhibits a rich tapestry of meaning that etymologically begins with signifying respect for educated men with social capital. It became a way of addressing bureaucrats who do clerical work during colonial times and increasingly took on the negative connotations of kafkaesque experiences with Indian bureaucracy.

Enrollment into Aadhaar is voluntary. It involves collection of biometric (ten fingerprints, two irises, and facial photograph) and demographic (name, age, gender, and address) data from residents, which is used to issue a unique 12-digit number to them (UIDAI 2010). At its surface, Yogita's story was about the challenges of navigating Indian bureaucracy without Aadhaar. Bureaucrats agreed with Yogita's claim that she was married. But neither she nor her husband were enrolled into Aadhaar. Without their Aadhaar numbers, bureaucrats claimed to be helpless, their computers will not let them document their marriage. Her point was simple, yet poignant: *digital interfaces designed to follow the rules of a bureaucracy act like bureaucrats themselves*. Resisting enrollment became a difficult task, as Aadhaar increasingly began to mediate state-citizen relations in India (Singh and Jackson 2021).

I have often returned to Yogita's story over the course of my broader research on everyday experiences and struggles of data subjects in living with data infrastructures. In their seminal work on data infrastructures, Star and Ruhleder (1996) have argued that data systems become infrastructures when they draw on and sink into the background of existing practices of distributed and collaborative work. For Aadhaar, this work centers on organizing government services. Issuing marriage certificates is an example of such services. A 2006 Supreme Court ruling made marriage registration compulsory in India to address a range of issues from reducing child marriage to protecting women's rights in cases of unlawful abandonment and domestic violence. Yogita and her husband made numerous visits to the marriage registration office, but without Aadhaar, they struggled to get their marriage registered. They even got a lawyer involved. After all, in September 2013, the Supreme Court had issued an interim order on the public interest litigations against Aadhaar that, "No person should suffer for not getting the Aadhaar card inspite [sic] of the fact that some authority had issued a circular making it mandatory" (Justice K.S. Puttaswamy and Anr. vs. Union of India and Ors. 2018, SC India:20). Aadhaar was not mandatory for registering marriage, and yet the computers at the marriage registration offices seemingly made it so. Ultimately, the bureaucrats resolved the couple's struggles by letting them key in dots instead of digits for Aadhaar numbers. Their struggle in getting their marriage registered points to emerging forms of social, material, and political work that people as data subjects must do to align with data systems, which are increasingly becoming the infrastructure that sustains their relationship with diverse organizations.

In Yogita's example, the work that the couple had to do to get their marriage registered *at* the digital interface itself was eventually quite simple: key in dots instead of numbers in a data field. However, this work *at* the digital interface hides the enormous amount of work that they had to do to negotiate with bureaucrats at the marriage registration office to provide them with this option. They figured out a way to work *at* the digital interface by negotiating how to work *around* it with bureaucrats as human interfaces. Yogita's example provides the foundation of my core argument that *just like data systems have a frontend and a backend demarcated by their digital interface, the work that data subjects must do to live with data systems also has a frontend and a backend demarcated by how their activities are organized in relation to the digital interface*. While the frontend of this work comes into conscious reflection when focusing on the nature of data subjects' activities *at the digital interface* of data systems (such as entering and correcting data and managing its circulation and interpretation), the backend has also slowly begun to inform scholarly inquiry in increasing attention to the nature of underlying activities *around the digital interface* that also condition how data subjects interact with data systems (for example, in negotiations with human interfaces to data systems).

Although this chapter begins with an example from India, living with data infrastructures is a global phenomenon. To illustrate this point, I am going to use examples from my research on Aadhaar-enabled identification practices in India and credit repair practices in the United States. In the following sections, I dive deeper into the analytic strategy of using the digital interface as an organizing principle to investigate the work that data subjects must do to live with data. In the first section, I argue that this work is not invisible per se; the frontend of this work is the engine of the data economy. In analyzing this work using the language of frontend and backend, I show how both forms of this work have

received considerable scholarly attention. In the second section, I focus explicitly on backend work of data subjects. I narrate field stories where it remains unaccounted for in the ongoing infrastructuring of data in every aspect of ordinary life and yet is a necessary condition for: (1) securing representation in data systems (taking an example from my research on Aadhaar); and (2) overcoming the challenges posed by data-driven decisions (drawing on a participant's story from a research study[6] in collaboration with Malte Ziewitz[7] on following the efforts of low-income individuals in Upstate New York to improve their subpar credit scores). While there are multiple examples from both these research studies that can exemplify the nature of backend work, I have chosen one from each to showcase the everydayness of this work across geographies. Finally, to conclude, I reflect on the uneven distribution of this work among data subjects in harmonizing their way of life with and securing desired outcomes for life chances through data.

Researching Data Subjecthood

I begin with a rather simplistic description of how data systems operate in analyzing the work of data subjects: data provided by people at digital interfaces of such systems and/or collected from various data sources as *inputs are processed through an algorithmic blackbox to produce outputs*, which are often automated and personalized data-driven decisions that people, in turn, as data subjects must contend with (Bowker 2013; Pasquale 2015). This work of contending with outputs either happens on the same interfaces of data systems and/or now also on different interfaces of other services that rely on these systems. *I consider soliciting support from human interfaces to these data systems who may input data on behalf of data subjects or make sense of outputs for them as backend to the frontend of digital interfaces where data subjects interact with data systems by themselves.* Data subjects often must negotiate a way out of their struggles with data systems with such human interfaces and at times, may also need to pay a fee (or a bribe) in exchange for their services. An analytic focus on digital interfaces, thus, allows differentiation between:

1. *Frontend work*: Work that happens *at* the digital interfaces of data systems. Given enough data infrastructure literacy[8] (J. Gray, Gerlitz, and Bounegru 2018) and investments in user experience design, people can work at digital interfaces by themselves without needing any support. The frontend includes, but is not limited to, (a) providing inputs through, and (b) interacting with services and, more broadly, organizing everyday life by using outputs of the interfaces of social media, search engines, on-demand service and e-commerce platforms, and e-government portals.
2. *Backend work*: Work that happens *around* the digital interfaces and sinks into the background of data systems. The backend includes, but (again) is not limited to, (a) gaining access to the interfaces and, thereby, claiming membership in digital services, (b) everyday struggles of securing affordances for their way of life through outputs, and (c) forms of affective labor in

[6] In this study, we accompanied single mothers, artists, gig workers, community college students, and even accountants in their attempts to understand and fix their credit scores over a period of 14 months from August 2018 to November 2019. Using a combination of interviews, monthly participant diaries, and diary-interviews, we pieced together the experiences of people grappling with subpar credit scores and their folk understanding of the workings of the credit scoring system (see, Ziewitz and Singh 2021).

[7] Alternatively, Ziewitz frames the work of living with data as contending with everyday challenges of *living in the shadow of a platform* as a continuously monitored and rated data subject. He elaborates on this using the example of search engine optimization as an industry operating in the shadow of search engines (Ziewitz 2017).

[8] The term "data infrastructure literacy" (J. Gray, Gerlitz, and Bounegru 2018) alludes to the need for investing in public understanding of working with data not just as a resource to organize practices (data skills), but also as the new condition for wider sociocultural changes in society (data politics, data culture, and data sociology).

> building a community around shared experiences of algorithmic harms. This work depends on the capacity of data subjects to figure out aspects of the system that are inaccessible and to create understandings, practices, workarounds as needed. It requires engaging with other data subjects and/or human interfaces and securing their support to get data systems to align with particularities of life circumstances. Calling customer support is a typical example where the support staff is expected to resolve struggles of data subjects as customers.

By calling out this difference, I do not mean to imply that the frontend is visible, while the backend is invisible. Rather, my objective is to showcase how the ability to work at digital interfaces and make sense of the workings of data systems is variably distributed and not equally accessible to all. It, thus, requires more effort from some data subjects than others to get the same work accomplished at digital interfaces. While these efforts come in diverse forms, they remain unpaid[9], but they must be performed as a precondition of everyday life in the data economy.

The frontend work has received considerable scholarly attention over the years. While mapping the expanse of this scholarship is outside the purview of this chapter, it often uses the interface as an analytic resource to make the work of data subjects visible. This scholarship spans across: (1) reliance of the data economy on "free services" in exchange for "free labor" of internet users (Terranova 2000) that, by extension, has created the conditions for economization of attention (Davenport and Beck 2001); (2) capturing the enthusiasm around crowdsourcing through notions of "wisdom of the crowds" (Surowiecki 2005), and critiquing the conditions of the production of labor that sustain the commercial web using the figure of the "digital housewife" (Jarrett 2015); (3) emergence of human data interaction (HDI) as a field of information science research that seeks to transform the passive collection of personal data in everyday interactions with data systems into an active reflection among data subjects on managing and controlling flows of their personal data (Crabtree and Mortier 2015); (4) disentangling the nature of contemporary participation in data systems (Kelty et al. 2015); and finally, (5) the rise of the sharing economy powered by on-demand service platforms (Sundararajan 2017).

In comparison, scholarly attention to backend work is only beginning to take shape. While there has been a consistent focus on algorithmic biases and arguments for improving fairness and accountability in operation of data systems (for example, Edelman 2011; Sweeney 2013; O'Neil 2016), explorations of the everyday struggles of data subjects is relatively new. Research in this domain has increasingly come to consider this work for reasons ranging from articulating agency and voice of data subjects from the bottom up (Couldry and Powell 2014) to formulating data-oriented notions of social justice to inform the work of data activists and practices of data-driven decision making (Taylor 2017; Kennedy 2018). Scholars have increasingly come to focus on: (1) how data subjects engage in mundane forms of sense-making to understand how algorithms work (Bucher 2017; Pink et al. 2017; Shen et al. 2021) and resist algorithmic power (Velkova and Kaun 2021); (2) how data subjects contend with surveillance, inequality, precarity, and discrimination perpetuated by data systems (Browne 2015; Eubanks 2017; Noble 2018); and finally (3) exploring everyday life and data subject communities in the gig economy (Raval 2020; Qadri 2020).

---

[9] This work performed by data subjects has deep resonance with Ivan Illich's notion of "shadow work" (1981). Illich used shadow work to conceptually address invisible, institutionally unrecognized, and unpaid work that is essential for the maintenance of its counterpart: visible and recognized waged work (Illich 1981; see also Star and Strauss 1999; Radin 2017). Illich further elaborates that, "To grasp the nature of shadow work we must avoid two confusions. It is not a subsistence activity; it feeds the formal economy, not social subsistence. Nor is it underpaid wage labor; its unpaid performance is the condition for wages to be paid. […] The creation of professionally supervised shadow work has become society's major business" (1981, 100–114). Most work that data subjects do in producing data for any digital service is unpaid. Supervising and facilitating this unpaid work is the business model for most digital services (see also, Jarrett 2015). As data systems pervade every aspect of ordinary life, this work, however, is increasingly also becoming subsistence activity for the poor and the systemically marginalized (Eubanks 2017).

Liminal to both frontend and backend work is equally important scholarship on the machinations of data extractivism and colonialism (Couldry and Mejias 2018; Ricaurte 2019), invisible forms of 'ghost' work that sustains the workings of the algorithmic blackbox (Irani and Silberman 2013; M. L. Gray and Suri 2019), and efforts to navigate seams of data systems (Singh and Jackson 2017) and resist oppression through data from the margins (Milan and Treré 2020). The digital interface, thus, is only a *partial solution* to the problem of analytically categorizing the diverse forms of work that data subjects do to live with data systems. However, it provides an analytic opening into backend work that does not get represented on, within, and through data systems but is crucial for their operation. In the next section, I explore two examples of this work where I focus on everyday interactions of data subjects with human interfaces to data systems. I focus on struggles over making the "right" inputs to secure access to a data system in the first example from India and explore how outputs are negotiated in pursuit of better life chances in the second example from the United States.

Forms of Backend Work

Harmonizing life with inputs

A crucial aspect of backend work is the struggle over securing access to a data system. Generally, gaining access[10] is as simple as registering (signing up) on a website. Registration often involves a set of mandatory questions that must be answered as the initial condition for access. It constitutes Know-Your-Customer/Citizen (KYC) requirements that are often preconditions of accessing any service. It is an initial resource for authenticating identity to authorize access and personalizing such services. The process of providing biometric and demographic data during Aadhaar enrollment exemplifies registration. For most Indian residents, this process has been fairly straightforward, involving visit to an enrollment agency where an enrollment operator inputs their data into the enrollment client software using a multi-lingual interface that displayed information in English and the most common local language of the region.[11] For others it produced a recursive set of challenges. These challenges were experienced by specific populations, for example: (1) struggles with lack of distinct biometric features among the elderly and manual laborers (Rao 2013; Khera 2018); (2) challenges of discrimination and troubles with certifying gender identity among the transgendered (Sethi 2012; M. Rajadhyaksha 2013); and (3) exclusion due to lack of a proof of identity and address among the homeless (A. Rajadhyaksha 2013). These struggles to secure a place in core data categories recursively create conditions for rendering individual identity and social history of data subject populations unaccountable (Star and Bowker 2007) as Aadhaar sinks into the background of everyday life in India.

Anahita, a researcher working on the role of Common Service Centers (CSCs)[12] in implementing eGovernance projects in India, brought a peculiar form of this recursion to my attention during a shoptalk phone call to discuss our respective fieldwork on Aadhaar:

> I was sitting in one of the CSCs in Rajasthan. I spoke to [Sana] about her experience of enrolling her son [Faizal] into the [Aadhaar] project quite a few times. […] They used to live in UP [Uttar Pradesh], then they moved to Rajasthan. She got her son

[10] Unpacking access must always begin with the work of overcoming the challenge of the digital divide in getting to the interface of data systems in the first place and the infrastructural constraints that perpetuate it. The process of becoming a data subject, however, begins after the barrier of digital divide is overcome.

[11] Data entry is displayed in English and Hindi in most parts of northern India; it is displayed in English and the regional language of the states, such as Tamil, Gujarati, Assamese etc. in other parts of the country.

[12] Common Service Centers (CSCs) are local offices created by the Indian government in rural and remote locations where access to internet is otherwise challenging for digital delivery of public services, such as paying electricity bills and checking the status of bureaucratic applications online.

> enrolled in UP in an [enrollment] camp. When she enrolled, they did not give her an enrollment receipt[13]. The operator told her that the receipt will be issued tomorrow, and she can come and pick it up then. She went the next day, and the camp was not there anymore. Her son's Aadhaar letter[14] never came. The rest of her family received their letters, but her son's letter never arrived and because she did not have an enrollment receipt, she could not follow it up.
>
> Then, they moved to Rajasthan in 2012 and here in Rajasthan, she got her son enrolled four times and still his Aadhaar card has not come. Now, she has enrollment receipt from all the four attempts. The operator looked it up and said that she had made quite a few mistakes. For example, her son's name is Faizal and some operator has entered the spelling of his name as 'Faisal' and others have written it as 'Faizal'. [...] The operator was scolding her that she should have checked this more thoroughly and how can she make multiple attempts like this. [If she had been literate, she would not have needed] the operator's help in resolving this problem.
>
> Now, the new problem is that her son has just passed class X and he needs to be admitted into class XI and the school is demanding his Aadhaar number to admit him in the next class. So, my confusion is that if biometrics are collected during enrollment to ensure uniqueness, how can her son be enrolled four times into the system (Anahita, Personal Communication, 25 July 2017).

Addressing Anahita's confusion requires a detour into the backend of Aadhaar as a data system. After data collection for enrollment is complete, the biometric data provided by an enrollee is checked against biometric data on all existing Aadhaar records. A new Aadhaar number is issued only after this deduplication process certifies that the enrollee's data is unique. Confusion around multiple enrollments stems from treating data collection and deduplication as synchronous processes. Deduplication, however, happens much later after the enrollment data packet is processed. During the data collection phase, enrollment data packets can be generated on an enrollee any number of times. If a person is already issued an Aadhaar number, these data packets are rejected as attempts to create a duplicate entry. In principle, while there can be multiple attempts at enrollment, there can only be one Aadhaar number generated for each enrollee.

In Sana's case, an Aadhaar number was probably already issued for her son, Faizal, although his Aadhaar letter was not delivered to them. In such cases, enrollees must work towards locating their Aadhaar record, if a number is generated, and then update it with new data. This backend work was challenging for several reasons: First, Aadhaar number is a prerequisite for updating an Aadhaar record. Faizal did not have his number or even know that he had one. Second, checking whether Aadhaar number was issued for him requires the enrollment receipt, but they did not have it from their first enrollment attempt. Third, figuring out what his Aadhaar number might be using his biometric features was not possible because authentication also requires an Aadhaar number. Aadhaar authentication involves a 1:1 comparison; biometric data provided is checked against the biometric data stored for an Aadhaar number. Enrollment is the only time when biometric data provided is compared against the entire database.

[13] This receipt is the bureaucratic proof of completing enrollment issued by the enrollment agency to enrollees.

[14] This letter is sent to enrollees with information about their Aadhaar number and demographic details stored in the Aadhaar database.

Re-enrollment would help only to the extent that it provides the operator with more data about why the enrollment attempt failed and indicate presence of an existing Aadhaar record. Instead, the operators were simply re-enrolling Faizal and inputting his data often incorrectly into the interface again and again with no success. The different forms of work, with uneven chances of success, performed by Sana, Faizal, and the specific populations who struggle with Aadhaar enrollment to harmonize their life situation with requisite inputs for enrolling into Aadhaar exemplify backend work. This work is a necessary condition for navigating everyday life as Aadhaar becomes the data infrastructure of certifying identity and completing KYC requirements in India. It is never explicitly recognized and becomes invisible after an Aadhaar number is secured by enrollees as data subjects. More importantly, it showcases that data records are not static representations of data subjects.

Contending with outputs

A data system is not the end; it is often a means to several ends. Such systems become the infrastructure for organizing services through their outputs such as lists, rankings, scores, recommendations, ratings, and more broadly, results of retrieval and sorting through data that make everyday life easier for some at the expense of others. They simultaneously exhibit function creep wherein a system designed for one purpose can be repurposed for another. A classic example of this dynamic is the use of credit scores in the United States, which was meant to operate as a loose index of an individual's creditworthiness (Lauer 2017). However, through processes of "off-label use," the scope of using these scores has moved far beyond financial credit as in cases of loans, mortgages, and insurance to underpinning forms of social credit in situations such as making decisions on hiring, renting, and even, online dating (Rona-Tas 2017). As the scope of using credit scores expands, so does the scale of the challenges faced by people who have subpar credit scores. Their struggles often begin at an individual level, credit scores after all represent an individual's creditworthiness. However, our research study on following credit repair journeys of participants with subpar credit scores highlighted that these struggles expand beyond the individual to become a family concern. A data subject's struggles with their data representation expands to encompass other family members. The family as a unit must contend with how it is represented through data.

A stark example of this insight emerged in conversations with one of our research participants, Maurice, who is a black professional artist with variable project-based income and engaged to Alexis, a firefighter. At the cusp of his married life, Maurice is proactive about improving his credit score, which was in the mid-650s.[15] During the 14-month long study, Alexis became pregnant, and they decided to rent a different place with more space to raise their child. However, this turned out to be a difficult ordeal because Alexis's credit score was in the mid-400s. As Maurice described it during a check-in with us:

> We were looking for an apartment. We were searching for something, um, [bigger]. We found something we really liked. It was a three-bedroom townhouse. And we went to fill out the application, we had gotten denied.
>
> It was the first conversation that we had… about just credit. And even with us being together for over two years… we had not had a direct conversation about that aspect of our finances (Maurice, check-in on diary entries, 25 February 2019).

[15] Credit score of the same person can be different across various bureaus and providers in the United States. Their range can be illustrated through broad descriptive categories such as "bad" (<600), "poor" (600–649), "fair" (650–699), and "good" (>700).

Talking about debt and money is often considered a taboo in the United States. Avoiding conversations about credit scores seems like a natural extension of this taboo, although this is changing with credit scores becoming a feature of online dating profiles. Certain life circumstances can also make such conversations unavoidable. Maurice had to have a conversation with Alexis because her score was too low for them to be able to rent their dream townhouse.

Rejection, however, was only the beginning of his efforts. He did not lose hope and made several attempts to meet the property manager:

> It was my third [attempt]. I went up there. I put on a shirt and tie, and … I said, "I would like to speak to the property manager. I know we got denied on this computer, but I wanna speak to a human."
>
> I said, "Listen. My fiancé loves this place. We have a baby on the way … and this is where she wants to be… What can we do to make this possible for her?"
>
> He asked me a few questions, and he said, "There is one loophole that we could do. I'm not sure if it'll work. I'm not sure where your credit is."
>
> I said, "Well, my credit is fine. It's not the best, but it's fine."
>
> So, he said, "There is a loophole. We'll put you in, and pending your approval, we'll list her as the roommate on the lease."
>
> And within five minutes, I had gotten approved (Maurice, check-in on diary entries, 25 February 2019).

Maurice was finally able to rent their dream townhouse. With credit score's orientation towards the individual, the data representation of their family's creditworthiness is malleable, given the support of the creditor. Maurice succeeded because he represented Alexis as his roommate on the lease. For the purposes of the property management system that evaluated rental applications, they were not a family. At times, success in contending with outputs is a matter of how data can be made to represent a data subject to achieve desired outcomes. This, of course, requires further ethical reflection on walking the fine line between legitimate participation in and illegitimate manipulation of data systems (Ziewitz 2019). Implicit in the ways in which data systems track behavior is the expectation that data subjects should behave as if these tools are not there (Zuboff 2019). There is, thus, an ongoing tension around the role that data subjects must be afforded in shaping their data. The smaller the role, the more oppressive data systems become. The greater the role, the easier it becomes to "game" them (Ziewitz 2019). In moments when data systems do not align with particular life circumstances of certain data subjects, their backend work[16] of getting these systems to work for them foregrounds this tension. This tension is at the heart of the mutual shaping of data subjects and data systems. More importantly, they highlight that data subjects are not passive recipients of data-driven decisions.

[16] Such backend work aligns with the notion of "repair work" required to make a data system work in an existing professional context—data systems often encounter breakages that must be smoothened over by repairing them (Elish and Watkins 2020). With backend work, however, I am implying work of aligning particularities of life with standardized data categories, which is easier for some data subjects and much more difficult for others.

Conclusion

This chapter uses digital interfaces as an organizing principle to analyze the unevenly distributed work that people as data subjects must do to live with data systems. I exemplify diverse forms of backend work using my research on everyday experiences of data subjects in India and the United States. My examples illustrate that *neither are data records static representations of data subjects, nor are data subjects passive recipients of data-driven decisions.* The mutual shaping of data records and everyday lives of data subjects happens not only at the interfaces of data systems, but also around them and often relies on conviviality of human interfaces. Sana did not receive their support, while Yogita and Maurice were eventually able to secure it. Yogita got a lawyer involved; Maurice put on a shirt and tie. They both projected social capital in their negotiations, while Sana reflected on not being literate as a barrier. The difference in the outcomes of their predicaments illustrates how social positioning plays a crucial role in overcoming the challenges of living with data.

More broadly, data-driven decision making is contingent on historical datasets and these datasets embed the societal values of the time of their construction (Bowker 2013; Busch 2014). The inputs to and the outputs based on these datasets do not represent the world we wish to live in; they represent existing biases of the world we already live in. The challenges raised by inability to secure representation through core data categories of a system and unfavorable outputs often mirror the uneven structural consequences of societal inequities along well-recognized intersections of gender, race, class, and caste. Diverse forms of backend work emerge in contending with such uneven challenges as data systems configure life chances in data economy. The need and the ability to do this work is unequally distributed among data subjects and is deeply intertwined with their position on these intersections. *Backend work is the uneven cost of living in a data-driven world.* On the one hand, the stories of Yogita and Maurice provide a silver lining to these unequal struggles. They showcase that data subjects can find creative and practical solutions to their struggles with a little help from intermediaries. These solutions often involve securing the support of human interfaces to data systems and working *around rather than with* data systems. On the other hand, Sana's story is a stark reminder that these solutions are not equally accessible to all. At times, it is not just the data system, but also the human interfaces that contribute to alienation of data subjects from the data economy.

This brings me to my final point. With increasing calls to focus on everyday experiences of data subjects (Kennedy 2018), it is crucial to attend to the role and intermediation work of human interfaces. A crucial aspect of backend work for data subjects is developing and leveraging relationships with humans in the loop to overcome their emerging and ongoing struggles with data systems. This work is a feature of a constitutive tension of organizing for data-driven services: Although one of the core purposes of deploying data systems is to remove intermediaries, they play a key role in making data systems work for or against data subjects in the last mile delivery of services. While frontline customer service operators are key actors in such negotiations, there are also increasing forms of communal knowledge sharing among data subjects about their experiences of making data systems work for them (DuFault and Schouten 2020; Kear 2021). Understanding how these data communities are formed and how they exemplify novel forms of backend work is crucial to our analytic efforts of unpacking everyday life in a data-driven world. These efforts must begin with the insight that data systems do not act alone; data-driven and human judgment are inevitably enmeshed in charting the ethical landscape of everyday life with data systems. After all, an interface as much human as it is digital.

References:

Aiyar, Shankkar. 2017. *Aadhaar: A Biometric History of India's 12-Digit Revolution.* Delhi: Westland.

Bowker, Geoffrey C. 2013. "Data Flakes: An Afterword to 'Raw Data' Is an Oxymoron." In *"Raw Data" Is an Oxymoron*, 167–171. Cambridge, Massachusetts: MIT Press.

Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Durham, NC: Duke University Press.

Bucher, Taina. 2017. "The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms." *Information, Communication & Society* 20 (1): 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Burrell, Jenna. 2015. "How the Machine 'Thinks:' Understanding Opacity in Machine Learning Algorithms." http://ssrn.com/abstract=2660674.

Busch, Lawrence. 2014. "Big Data, Big Questions| A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large Scale Data Sets." *International Journal of Communication* 8 (1727–1744). http://ijoc.org/index.php/ijoc/article/view/2160/1160.

Couldry, Nick, and Ulises A Mejias. 2018. "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject." *Television & New Media* 20 (4): 336–49. https://doi.org/10.1177/1527476418796632.

Couldry, Nick, and Alison Powell. 2014. "Big Data from the Bottom Up." *Big Data & Society* 1 (2): 1–5. https://doi.org/10.1177/2053951714539277.

Couldry, Nick, and Jun Yu. 2018. "Deconstructing Datafication's Brave New World." *New Media & Society*, May, 1461444818775968. https://doi.org/10.1177/1461444818775968.

Crabtree, Andy, and Richard Mortier. 2015. "Human Data Interaction: Historical Lessons from Social Studies and CSCW BT - ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway." In , edited by Nina Boulus-Rødje, Gunnar Ellingsen, Tone Bratteteig, Margunn Aanestad, and Pernille Bjørn, 3–21. Cham: Springer International Publishing.

Davenport, Thomas H., and John C. Beck. 2001. *The Attention Economy: Understanding the New Currency of Business*. Boston: Harvard Business Press.

DuFault, Beth Leavenworth, and John W Schouten. 2020. "Self-Quantification and the Datapreneurial Consumer Identity." *Consumption Markets & Culture* 23 (3): 290–316. https://doi.org/10.1080/10253866.2018.1519489.

Edelman, Benjamin. 2011. "Bias in Search Results: Diagnosis and Response." *Indian JL & Tech.* 7: 16–32. http://www.ijlt.in/archive/volume7/2_Edelman.pdf.

Elish, Madeleine Clare, and Elizabeth Anne Watkins. 2020. "Repairing Innovation: A Study of Integrating AI in Clinical Care." New York: Data & Society Research Institute. https://datasociety.net/wp-content/uploads/2020/09/Repairing-Innovation-DataSociety-20200930-1.pdf.

Eubanks, Virginia. 2017. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Gray, Jonathan, Carolin Gerlitz, and Liliana Bounegru. 2018. "Data Infrastructure Literacy." *Big Data & Society* 5 (2): 1–13. https://doi.org/10.1177/2053951718786316.

Gray, Mary L., and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston and New York: Houghton Miffin Harcourt.

Hyysalo, Sampsa, Torben Elgaard Jensen, and Nelly Oudshoorn, eds. 2016. *The New Production of Users: Changing Innovation Collectives and Involvement Strategies*. New York and London: Routledge.

Illich, Ivan. 1981. *Shadow Work*. Boston and London: Marion Boyars.

Irani, Lilly C, and M Six Silberman. 2013. "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 611–20. CHI '13. New York, NY, USA: ACM. https://doi.org/10.1145/2470654.2470742.

Jarrett, Kylie. 2015. *Feminism, Labour and Digital Media: The Digital Housewife*. New York and London: Routledge.

Justice K.S. Puttaswamy and Anr. vs. Union of India and Ors. 2018. Writ Petition (Civil) No. 494 of 2012, SC India.

Kear, Mark. 2021. "The Moral Economy of the Algorithmic Crowd: Possessive Collectivisim and Techno-Economic Rentiership." *Competition & Change*, February, 1024529421990496. https://doi.org/10.1177/1024529421990496.

Kelty, Christopher M., Aaron Panofsky, Morgan Currie, Roderic Crooks, Seth Erickson, Patricia Garcia, Michael Wartenbe, and Stacy Wood. 2015. "Seven Dimensions of Contemporary Participation Disentangled." *Journal of the Association for Information Science and Technology* 66 (3): 474–88. https://doi.org/10.1002/asi.23202.

Kennedy, Helen. 2018. "Living with Data: Aligning Data Studies and Data Activism through a Focus on Everyday Experiences of Datafication." *Krisis: Journal for Contemporary Philosophy*, no. 1: 18–30. https://krisis.eu/living-with-data/.

Khera, Reetika, ed. 2018. *Dissent on Aadhaar: Big Data Meets Big Brother*. Delhi: Orient BlackSwan.

Lauer, Josh. 2017. *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*. New York: Columbia University Press.

Marcus, George E. 1995. "Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography." *Annual Review of Anthropology* 24 (1): 95–117. https://doi.org/10.1146/annurev.an.24.100195.000523.

Milan, Stefania, and Emiliano Treré. 2020. "Big Data from the South(s): An Analytical Matrix to Investigate Data at the Margins." In *The Oxford Handbook of Sociology and Digital Media*, edited by Deana A. Rohlinger and Sarah Sobieraj. Oxford: Oxford University Press.

Nilekani, Nandan, and Viral Shah. 2015. *Rebooting India: Realizing a Billion Aspirations*. Gurgaon: Penguin Books India.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Oudshoorn, Nelly, and Trevor Pinch, eds. 2003. *How Users Matter: The Co-Construction of Users and Technology (Inside Technology Series)*. Cambridge and London: MIT Press.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA and London: Harvard University Press.

Pink, Sarah, Shanti Sumartojo, Deborah Lupton, and Christine Heyes La Bond. 2017. "Mundane Data: The Routines, Contingencies and Accomplishments of Digital Living." *Big Data & Society* 4 (1): 1–12. https://doi.org/10.1177/2053951717700924.

Qadri, Rida. 2020. "Algorithmized but Not Atomized? How Digital Platforms Engender New Forms of Worker Solidarity in Jakarta." In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 144. AIES '20. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3375627.3375816.

Radin, Joanna. 2017. "'Digital Natives': How Medical and Indigenous Histories Matter for Big Data." *Osiris* 32 (1): 43–64. https://doi.org/10.1086/693853.

Rajadhyaksha, Ashish, ed. 2013. *In the Wake of Aadhaar: The Digital Ecosystem of Governance in India*. Bengaluru: Centre for the Study of Culture & Society. https://pad.ma/documents/PI.

Rajadhyaksha, Madhavi. 2013. "Will India Recognize a Third Gender?" *Times of India*, December 8, 2013. http://timesofindia.indiatimes.com/india/Will-India-recognize-a-third-gender/articleshow/27050953.cms.

Rao, Ursula. 2013. "Biometric Marginality." *Economic and Political Weekly* 48 (13): 72–77. http://www.epw.in/review-urban-affairs/biometric-marginality.html.

Raval, Noopur. 2020. "Platform-Living: Theorizing Life, Work, and Ethical Living after the Gig Economy." UC Irvine. https://escholarship.org/uc/item/4qc0x3mw.

Ricaurte, Paola. 2019. "Data Epistemologies, The Coloniality of Power, and Resistance." *Television & New Media* 20 (4): 350–65. https://doi.org/10.1177/1527476419831640.

Rona-Tas, Akos. 2017. "The Off-Label Use of Consumer Credit Ratings." *Historical Social Research / Historische Sozialforschung* 42 (1 (159)): 52–76. http://www.jstor.org/stable/44176024.

Sethi, Ashpreet. 2012. "Getting Aadhaar Card Big Challenge for Transgenders." *Deccan Herald*, May 17, 2012. http://www.deccanherald.com/content/250353/getting-aadhaar-card-big-challenge.html.

Shen, Hong, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. "Everyday Algorithm Auditing: Understanding the Power of Everyday Users in Surfacing Harmful Algorithmic Behaviors." *Proc. ACM Hum.-Comput. Interact.* 5 (CSCW2). https://doi.org/10.1145/3479577.

Simone, AbdouMaliq. 2004. "People as Infrastructure: Intersecting Fragments in Johannesburg." *Public Culture* 16 (3): 407–29.

Singh, Ranjit, and Steven J. Jackson. 2017. "From Margins to Seams: Imbrication, Inclusion, and Torque in the Aadhaar Identification Project." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, CO: ACM.

———. 2021. "Seeing Like an Infrastructure: Low-Resolution Citizens in the Aadhaar Identification Project." *Proceedings of the ACM on Human-Computer Interaction* 5 (CSCW2): 315:1-315:28. https://doi.org/10.1145/3476056.

Star, Susan Leigh, and Geoffrey C Bowker. 2007. “Enacting Silence: Residual Categories As a Challenge for Ethics, Information Systems, and Communication.” *Ethics and Information Technology* 9 (4): 273–80.

Star, Susan Leigh, and Karen Ruhleder. 1996. “Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces.” *Information Systems Research* 7 (1): 111–34.

Star, Susan Leigh, and Anselm Strauss. 1999. “Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work.” *Computer Supported Cooperative Work (CSCW)* 8 (1–2): 9–30.

Sundararajan, Arun. 2017. *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. Cambridge, MA: MIT Press.

Surowiecki, James. 2005. *The Wisdom of Crowds*. New York: Anchor Books. http://books.google.com/books?id=Jrhsf5WlBxMC&pgis=1.

Sweeney, Latanya. 2013. “Discrimination in Online Ad Delivery.” *Commun. ACM* 56 (5): 44–54. https://doi.org/10.1145/2447976.2447990.

Taylor, Linnet. 2017. “What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally.” *Big Data & Society* 4 (2): 1–14. https://doi.org/10.1177/2053951717736335.

Terranova, Tiziana. 2000. “Free Labor: Producing Culture for the Digital Economy.” *Social Text* 18 (2): 33–58. http://www.electronicbookreview.com/thread/technocapitalism/voluntary?mode=print.

UIDAI. 2010. *UIDAI Strategy Overview: Creating a Unique Identity Number for Every Resident in India*. New Delhi: Unique Identification Authority of India. https://www.dropbox.com/s/eg9p5uzucsd9t5r/UIDAI_Strategy_Overview_2010.pdf?dl=0.

Velkova, Julia, and Anne Kaun. 2021. “Algorithmic Resistance: Media Practices and the Politics of Repair.” *Information, Communication & Society* 24 (4): 523–40. https://doi.org/10.1080/1369118X.2019.1657162.

Woolgar, Steve. 1990. “Configuring the User: The Case of Usability Trials.” *The Sociological Review* 38 (S1): 58–99. https://doi.org/10.1111/j.1467-954X.1990.tb03349.x.

Ziewitz, Malte. 2017. “Shady Cultures.” Fieldsights. 2017. https://culanth.org/fieldsights/shady-cultures.

———. 2019. “Rethinking Gaming: The Ethical Work of Optimization in Web Search Engines.” *Social Studies of Science* 49 (5): 707–31. https://doi.org/10.1177/0306312719865607.

Ziewitz, Malte, and Ranjit Singh. 2021. “Critical Companionship: Some Sensibilities for Studying the Lived Experience of Data Subjects.” *Big Data & Society* OnlineFirs: 1–13. https://doi.org/0.1177/20539517211061122.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Hachette Book Group.