Introduction

DATA LIVES ON

Ana V. Sanchez is dead. Writing these words floods me with many emotions: sadness, anger, hopelessness. Above all, I am filled with shame. Ana and I spoke perhaps three times before she died of brain cancer several years ago. Ana was not my cousin or sister, or even a very close friend. She was more of an acquaintance, a vaguely familiar face that I would sometimes see in a crowd at a party.

"We met at that art opening in 2009, right?" Not much else connected Ana and me. The last time we saw each other, she had to ask me to remind her of my name. It saddened me when she asked because I detected a slight slur in her question. "Mary," I reminded her. And she beamed. "Maria? Oh, yes, yes, Mary."

I first met Ana in 2009, at an art opening, where we instantly hit it off. Ana, who had a piece in the show, was a fiber artist, and much of her work involved large installations—of luxuriant red satin sewn into hanging walls of peony petals, or of cotton balls stitched together with fine red thread to make clouds that seemed to bleed. In her hands, lifeless fabric was transformed into verdant invocations of life. Thinking back to the first time I met Ana, I recall her face radiating an inner warmth and kindness. Her chestnut hair was long and glossy, which seemed to mimic the walls of cascading satin petals. As we circulated in the gallery to look at each artist's work, we talked. I learned that Ana had grown up in Spain, and that while some of her relatives, like her dad, lived in Puerto Rico, she was by herself in Philadelphia, far away from her family. I assumed that Ana found herself, as I did, living in a society that valorizes individualism and self-reliance, and that tends to see personal hardship, especially in regard to health or debt, as a personal failing, rather than as a brutal symptom of systemic dispossession and injustice.

The second time I saw Ana, it was at the opening of a group show. This time Ana was wearing a white shirt with a little black bow tie clipped onto the collar, black trousers, and black shoes. Her lush hair was pinned back into a neat bun. Her art-making alone could not support her financially, and her catering jobs were often for art openings that featured the work of her friends and colleagues in the tight-knit art community in Philadelphia. Like so many American workers, she learned to live with contingent contracts, precarious income, and little stability. Ana learned to survive, barely, with no health insurance. We only saw each other briefly that night, when she offered me an hors d'oeuvre, a miniature quiche, from her tray as an excuse to say hello and talk for a few minutes.

The last time I saw Ana, it was after her third surgery to remove the brain tumor, at yet another art opening. This time, she was just visiting, not serving food or drinks. Her hair was now brittle and completely gray. It was still long, which enabled Ana to pin it over the side of her head where a five-inch scar snaked around her skull. Ana saw me, and in an instant, tears welled up in her eyes. She had a catch in her throat as she thanked me for my donation to her YouCaring site, an online crowd-sourced fundraising campaign for medical debt. I had given fifty dollars about a year earlier to help pay for one of her brain surgeries. As far as I knew, the hospital that was treating her cancer hadn't helped her to set up the online campaign, which is one of several third-party financing services that medical practices now offer insured, underinsured, and uninsured patients to help them pay for their treatment costs. That evening was the last time I saw Ana alive. She died later that year, in November, just a few months shy of her forty-first birthday.

Cancer may be among the costliest medical conditions to treat, and its treatment places a huge financial strain on the individual patient, their families, and the healthcare system broadly (Mariotto et al. 2011; Park and Look 2019). Even for those who have insurance, the complex billing processes and co-pays—costs associated with procedures and drugs that insurance does not cover—and other out-of-pocket expenses force many into bankruptcy. This treatment-related devastation has been named "financial toxicity" among cancer healthcare researchers, and is related to lower-quality care, poorer outcomes, and higher mortality rates (Zafar 2016; Zafar and Abernethy 2013). The financial toxicity associated with oncological medical treatments makes patients sicker and less likely to survive. In consideration of these overwhelming financial burdens for patients and their families, but also to ensure that medical providers are paid, many hospitals and cancer treatment clinics now offer financial counseling. This counseling can include advising patients to look to alternative treatment financing, and helping to promote a patient's crowdfunding campaign to pay for cancer treatment (Berliner and Kenworthy 2017; Cohen et al. 2019). These days, too, when a patient's account goes into arrears with a medical provider, hospitals turn to debt collectors, and to the courts, to go after their patients for these unpaid debts. These patient debtors may even end up in prison.[1]

Since I didn't know Ana that well, I could only speculate as to the financial measures she had already exhausted, from maxing out her credit cards to borrowing from friends and family. I imagined that Ana's medical debt was considerable if she had turned to online crowdfunding to pay for her treatment. Ana's debt—information that would have embodied, and been marked by, the traumas and fears that accompany such financial liabilities—was no doubt fed into credit-risk algorithms as simple points of data, numbers that would have tanked her credit score (Dean et al. 2018; Zeldin and Rukavin 2007).[2] Before that, though, as a gig worker with no insurance, she had most likely opted out of earlier medical care that might have slowed the progression of her illness, or at the very least, improved the quality of her life.

The last time I saw Ana, all of the unbearable alienation, desperation, and shame brought on by our profit-driven healthcare system condensed into a hardness in our throats. We shared a common subjectivity; we were both subjects in a medical system driven by debt and data, where healthcare is not a right but a privilege.[3]

The shame that I felt that last night that I saw Ana continues to haunt me. I had given her fifty dollars: that's it, next to nothing in the grand scale of funding terminal brain cancer treatment with or without health insurance. I did not give her any of my time or care while she was dying. I didn't bother to write my representatives to demand healthcare as a human right nor did I even join a coalition, such as the Poor People's Campaign or groups advocating for Medicare for All. I couldn't manage to attend even one protest, or a hearing for healthcare financing reform. This angry shame that I carry boils down to this: Ana's gratitude for the charity of strangers, even for a measly fifty dollars delivered by the click of a button, was misplaced in that she was grateful for something that should have been hers by right. Ana died despite that tenuous shred of empathy shared with her, because in the brutal system of Hobbesian healthcare, there is little room for tenderness.

The United States has one of the most expensive healthcare systems in the world. Americans spend close to 18 percent of the gross domestic product (GDP, or total economic output) on healthcare (Centers for Medicare and Medicaid Services n.d.). Because it is technologically advanced and uses cutting-edge innovations such as machine learning and artificial intelligence (AI) diagnostics or the latest gene therapies, this high-cost healthcare system is cited as the best in the world. However, the most expensive does not mean the best, or even simply better, especially when it comes to equity. Many scholars who study healthcare policy in the United States, from healthcare economists to public health researchers, have long noted that these expenditures indicate that our healthcare system not only is inequitable, but also drives inequality (Nunn, Parsons, and Shambaugh 2020). The high costs in American healthcare result not from technological innovations or overutilization in general, but rather from the high prices of services and administrative expenses associated with a predominately for-profit system (Papanicolas, Woskie, and Jha 2018).

The cruelty of Ana's last few months struggling to survive brain cancer belies all the promises made about the high quality and effectiveness of the American healthcare system. These promises continue to go unfulfilled for millions of Anas in a society where shame and gratitude are the currencies transacted among us, instead of mutual commitment and obligation to one another. Why is it that in late capitalism, in for-profit healthcare, the means of exchange always place us at the sharp end of a transactional relationship, one that pits our survival against giving up control over our data and our financial futures? Medical capitalism forces many patients, insured and uninsured alike, into debt, pushing life-saving healthcare out of reach for more Americans and amplifying health and economic disparities. The COVID-19 pandemic has further elucidated and deepened these stark systemic inequities.

Ana passed away in November 2016. Just a few years later, the United States weathered another economic crisis and recession, brought on by a pandemic, as well as an uprising against systemic racism and violence. Unequal social conditions persisting from the 2008 global economic recession worsened, and many Americans remained in danger of getting sick or dying of the coronavirus; or being denied access to care, including the vaccine; or losing their savings, jobs, and homes. While the social inequities in health and wealth seem even worse than they were just a few years earlier, when Ana died, some things have not changed much. The skewed bargains that ensure that corporate interests control and profit from our data in healthcare and in finance persist and flourish. The tech giants, like Amazon and Google, as well as the health insurers and credit-reporting bureaus all reported record-breaking profits during the pandemic (Plott, Kachalia, and Sharfstein 2020).[4]

AFTER DEATH

I always saw a connection between what I imagined happened to Ana—and the possible ways that data about her oncological care and medical debt lived on after she died—and my own story of personal grief. At the time that Ana died, I was struggling with data concerning my own health and medical debt. These data were animated and given a rather phantasmagoric afterlife. For more than five years, a phantom haunted me. Like many ghostly apparitions, my phantom didn't have a body, but I sensed its presence in my day-to-day life. It didn't have a voice, but it spoke just the same. Though it was a complete stranger, my specter was uncannily familiar. It was intimately connected to my body, my health, and my life. The phantom grew to become my kin, my descendant: It was my data revenant. It was a ghostly creature whose body was made from the health, financial, and consumer data that lived on my credit cards in the form of medical debt or in the opaque information that comprised my credit score, in the cookies and browser history on my laptop, and in countless other digital crumb trails that I have left over the years.

You could say that the genesis of my data revenant, as well as this book, began on March 14, 2011—the day I learned that I had had my final miscarriage. That morning my doctor showed me that the cluster of glowing pixels on the ultrasound's monitor, which had vibrated with my baby's heartbeat only a week earlier, had gone still and silent. It was in the afternoon of that gray, early spring day that I met my data baby for the first time. Just like a shapeshifter or, as I came to understand it, a revenant, my data baby took on the shape of my fleshy baby that might have been. Instead of blood, skin, and bones, my data baby was composed of the free samples of baby things and other targeted marketing appeals sent directly into my mailbox or my email inbox for five years after my miscarriage. Where my pregnancy resulted in a miscarriage, my data baby gestated for nine months and was born at term, with no complications and completely healthy. Over those years, my data baby grew from being a newborn, to a toddler, to a preschooler, and finally, the last time that I heard from my data baby, it was getting ready to start kindergarten.

In the depths of my grief back then, I could not have imagined or known how the loss of my baby would eventually lead me to dedicate so much time, emotional labor, and psychic energy to studying something that most of us can't see or touch. At times, we can sense its presence humming in the background of our daily lives, and for some of us, its existence can make itself known, often in disruptive or harmful ways.

What we sensed happening "behind the scenes" has come under the glare of public scrutiny and concern. The *New York Times Magazine*'s 2012 investigative piece "How Companies Learn Your Secrets" uncovered how retailers like Target customize price markdowns to individual consumers by collecting data on in-store and online sales transactions, which they then run through predictive analytics. At several antitrust hearings, members of Congress publicly interrogated Big Tech executives—Alphabet (parent company of Google) CEO Sundar Pichai, Facebook executives Mark Zuckerberg and Sheryl Sandberg, and Twitter CEO Jack Dorsey—

on their industry's mishandling of millions of users' personal information and their platforms' role in generating dangerous pandemic and election misinformation and extremism (Ebeling 2016).[5] For many people data means not revolutionary promises of empowerment and liberation, but surveillance, loss of control and freedom, intrusions, division, and harm. The incremental and skewed bargains that we make concerning our data have detrimental consequences for our health, our finances, and our lives (Pasquale 2015; Zuboff 2019). Data about our lives, in all their varied forms, produced by us, lead existences that run parallel to our own "real life." The data, which often define us in ways that we struggle to recognize, can never represent the fullness of our lives. As I show, these data survive even death.

My research into all the places where our health and financial data live on started in 2014. Spurred on by my grief, anger, and the trauma of being haunted by data, I was determined to hunt down the pieces of my data baby in the disparate databases of marketers and reassemble them to make them whole. This research has defined my scholarship ever since. Through my investigations into how my health data ended up in the hands of marketers, I learned that the data baby was conceived with the purloined data taken from interactions in my day-to-day life. Data brokers harvested these data in places not limited to but including the pharmacy where I purchased over-the-counter ovulation test kits and the prescribed drugs used in in-vitro fertilization protocols, the fertility clinic where my credit card transactions paid for the copays and costs of medical treatment that my health insurance did not cover, and my computer and my phone where metadata about my online behaviors was generated as I visited websites about what I could expect in early pregnancy. I learned that my data went on to have multiple afterlives in all sorts of databases, owned and maintained by companies that I had never heard of, much less did business with. Some of my data were used to inform clinical trials or to boost the published success rates for the fertility clinic (because I did get pregnant, after all), and other bits of my data were segmented into marketing categories and commodified by data brokers, such as Experian plc. One of the largest data brokers, marketing services, and credit bureaus in the United States, Experian is as powerful as Alphabet when it comes to accessing and monetizing our data. However, most Americans know Experian only as a company that produces credit reports and credit scores. In recognition that the handling, trading, and monetizing of private health and financial information by powerful companies like Experian is widespread, in 2000, Congress amended parts of the Gramm-Leach-Bliley Act of 1999 (GLBA) that directly address the privacy and security of personally identifiable information. The legislation originally deregulated the financial services, insurance, and banking sectors to allow for mergers among these sectors; the amendments added privacy safeguards in an effort to protect consumers and patients from harmful cross-sector data exposures.[6] These legislative guardrails, however, did not prevent the consolidation of power over personal, private data into the hands of a few corporations in the financial services and credit information reporting sectors. Because these powerful players control vast amounts of our data, they can also determine outcomes in the high-stakes decision-making processes that use our data in healthcare and personal credit and banking services.

I wanted to find out how a company that I did not have any known direct contact with got hold of my private health data, especially data that, at the time, I naively considered to be mine, or at the very least, co-owned by myself and my doctor. I could not understand how a credit bureau could access information about my health, especially clinical data that most of us believe are protected by special privacy regulations; in particular, by the 1996 Health Insurance Portability and Accountability Act (HIPAA). I interviewed data brokers, data marketers, healthcare professionals (doctors, nurses, surgeons), bioinformaticians, pharmaceutical sales reps, and other patients who were haunted by their own data revenants. Through this process I uncovered how our data, especially information that is interpreted as health data, are harvested, commodified, and made to produce value for those that come to own our data. Taking up the role of a hardboiled detective, I wrote an autoethnographic noir about my investigations into data-marketing surveillance, the biodata economy, and how our medical data come to haunt us, in *Healthcare and Big Data: Digital Specters and Phantom Objects* (Ebeling 2016). With *Afterlives of Data,* I continue and expand upon the research started in my first book. Here my aim is to understand how our traumas, losses, and shame haunt our data, and how we are subjects of the double regimes of debt and data. Those subjectivities of trauma and loss, of debt and dispossession, become embodied as the data assets that produce value for the data economy at large. Often, that value does not return to us, the subjects of data and debt.

A NOTE ON METHODS

At a field site that I was researching for this book, one of my interlocutors and collaborators, Anmei, a nursing informatics researcher, aired her frustrations with a statistical model. She couldn't make the data fit the model, a conundrum I would often hear about from data scientists, and if she forced it, the model would no longer work. "Give me a number, a value, and it is either yes or no, and I can get to the truth," she said. "Qualitative research, sociology, what you do, is too difficult; the truth is not just one thing." Anmei's desire to stand outside the phenomenon under study by objectifying it through numbers is an impossibility for many sociologists, especially those who use qualitative methodologies like ethnography, my preferred method, to understand social phenomena.

In his book *After Method: Mess in Social Science Research* (2004), John Law warned that while "the social" appears to be definite, something discoverable that can be validated scientifically, social realities are fragmentary, multiple, complex, and always locally situated. The trick is to remember that the very social science methods developed to "know" these phenomena simultaneously help to create the same social realities that are being studied. Law noted that amid the ongoing debates about the nature of social reality, "the task is to imagine methods [that] no longer seek the definite, the repeatable, the more or less stable" (Law 2004, 6). Law did not propose a relativist vision of qualitative social research methods, but rather suggested that different practices will inevitably produce different study objects, different things, and that these objects shouldn't necessarily cohere or be reproducible. The findings that qualitative methods, especially ethnographic ones, produce are complex, messy, and fragile. This is true in the case of my own attempts at understanding and describing the complexity and the multiplicity of health and credit data. As one data marketer told me, all data can be health data.

I take this approach to underline the situatedness of sociological knowledge, as well as my own positionality and my relationships to social phenomena, and to emphasize the resemblance of data to a black box. Data's promise of total knowledge is dangerous and false. Much like sociological research, data produce their own reflexive, translational objects (Star and Griesemer 1989). So deeply embedded are data, and their uses by unseen others, in our day-to-day lives, we can never fully understand the processes by which data are harvested, decontextualized, and deployed to become the interpretable anchors for evidence or definitive factors that can predict the future. I have found that many of the practitioners, data brokers and analysts, themselves do not understand these transformations. Many of the experts that I interviewed for this book don't fully comprehend the density or breadth of the networks over which data traverse, nor can they break through the opacity of the algorithms and models deployed to interpret data. So how could I, as an outsider, do better than the professionals who work with health and credit data day in and day out? We are all insiders and we are all subjects, including the informaticians and practitioners I observed and interviewed, in the data-based society.

In the following chapters, I examine the equally disconcerting trends of the private sector's wholesale collection of consumer and financial data from a variety of sources, and the selling of this segmented information as health data to third parties. Over the years this research has taken me to hospitals, healthcare facilities, and private clinics where I interviewed medical practitioners and informatics researchers who handle patient data on a daily basis. I spoke to these groups to understand how private health data moves from the doctor's office into other sectors outside of medicine—with or without our consent. Once we are dispossessed of our data, we have little control over where they travel, how they are used, and whether our data will be used to grant us access to or deny us healthcare.

In these six years of conducting ethnographic, qualitative research, I have attended database marketing conferences and trade shows to speak with database marketers, data brokers, and sector analysts. Collecting information from these subjects helped me better understand how health and financial data become assets that are used to generate value for the credit-reporting and healthcare sectors. I have participated in patient privacy summits and academic symposia where I met with legal experts on privacy and clinical data. I have interviewed financial professionals who handle data commodities, asking them how health data is collected from credit card, retail banking, and other financial transactions. I entered all these field sites to learn more about how those working in healthcare generally, and in health informatics specifically, conceive of data as both the subject and product of profit-driven medicine.

Much of *Afterlives of Data* is the result of recent ethnographic fieldwork conducted at a university-based medical school's health informatics research institute; the fieldwork for my previous book also informs some of the research described here. The recent portion of the fieldwork occurred over a three-year period, starting in 2017 and going through the early days of the coronavirus pandemic in spring 2020. During my summers and semester breaks from the university where I teach in Philadelphia, I traveled to the medical school and hospital system, based in a Midwestern city, to learn more about how the data of tens of thousands of patients that the hospital serves are collected and operationalized into use cases, predictive scores, machine learning, and other algorithmic models in order to rationalize or to prove the efficacy of care, among many other purposes. My ethnographic fieldwork was doubly nested within a population health data research laboratory led by an epidemiologist and data scientist, whom I call Theresa. It was mostly in Theresa's lab where I learned more about how patient clinical information becomes part of massive databases maintained by healthcare providers and public and private insurers, or payers—often called data "lakes" and "oceans" by health informaticians. These databases are then trawled to make models and predictive data-based instruments designed to improve health outcomes for patients or, just as often, to create efficiencies for hospital systems under the pressures of value-based care.

Throughout this book, I have used pseudonyms to refer to the people that I interviewed or observed since commencing research in 2014. I have used pseudonyms, as well, for the institutions that I spent time in conducting fieldwork. The exceptions are the large data-marketing and health data conventions, such as Academy Health's Health Datapalooza or the Data Marketing Association's meetings and conventions, which retained their actual names. I refrain from using the name of the city where the university hospital system is located, other than to place it regionally in the Midwest of the United States.