and the institutions and machines that are meant to know it in their stead, between the practical capabilities of data technologies and the wider fantasies that give them legitimacy. In each one, we find troubling asymmetries in how different bodies are treated to different kinds of factmaking. If data expands the vistas of human action and judgment, it also obscures them, leaving human subjects to work ever harder to remain legible and legitimate to the machines whose judgment they cannot understand. Caught in an expanding and consolidating data market, we cannot simply seek more and better knowledge but must rethink the basic virtues and assumptions embedded in that very word. What kind of good does knowing do? Or, rather, what must our knowledge look like that it may do good? And who are we, with what kinds of capabilities and responsibilities, with what role to play in a data-driven society? As the truth of who we are and what is good for us is increasingly taken outside ourselves and human experience, the figure of the human subject—which, Foucault had warned, is a young and temporary thing[15]—is flickering uncertainly, unsure of the agency and moral responsibility we had worked so hard to attach to it.

1

Honeymoon Objectivity

In 2014, a baby-faced, twenty-two-year-old entrepreneur named James Proud crowdfunded a sleep-tracking device that promised to automatically monitor sleep patterns, provide a numerical score, and make recommendations for sleep behavior. That such functions were already available did not escape Proud. Beddit, a sleep sensor that we will revisit in chapter 4, had been crowdfunded a year before and already released to its backers. In response, Proud chose to emphasize his device's "simple, uncomplicated and useful" qualities; designed as a slick, minimalist off-white orb, it would merge invisibly into the everyday flow of attention and reflection. "We believe technology needs to disappear," said Proud; "everything in [our device] is just designed to fade away."[1] It would carry an equally simple and no-brainer name: Sense.

In 2017, James Proud, now twenty-five, announced the end of Sense.[2] Panned by some tech reviewers as a "fundamentally useless" object[3] and a glorified alarm clock, the device never quite delivered the quiet transformation of everyday life that its creator aspired to. Fundamentally, it proved not very good at making sense of human sleep. Users reported that any deviation from the presumed sleep scenario—for instance, a pet snuggling up in bed—would throw the device off entirely. The chaos of everyday life rarely conformed to the expectations of the tracking machine, even as its selling point was that it would discover truths about us that we cannot perceive ourselves. As Proud's team wound down operations, users began to report that their Senses were losing functionality. The orbs went mute and deaf to the data around them, a small monument to the unfulfilled promises of new technologies.

Technologies of datafication reconfigure what counts as truth and who—or what—has the right to produce it, and not simply through the success of indisputably superior machines or even their mundane

and the institutions and machines that are meant to know it in their stead, between the practical capabilities of data technologies and the wider fantasies that give them legitimacy. In each one, we find troubling asymmetries in how different bodies are treated to different kinds of factmaking. If data expands the vistas of human action and judgment, it also obscures them, leaving human subjects to work ever harder to remain legible and legitimate to the machines whose judgment they cannot understand. Caught in an expanding and consolidating data market, we cannot simply seek more and better knowledge but must rethink the basic virtues and assumptions embedded in that very word. What kind of good does knowing do? Or, rather, what must our knowledge look like that it may do good? And who are we, with what kinds of capabilities and responsibilities, with what role to play in a data-driven society? As the truth of who we are and what is good for us is increasingly taken outside ourselves and human experience, the figure of the human subject—which, Foucault had warned, is a young and temporary thing[15]—is flickering uncertainly, unsure of the agency and moral responsibility we had worked so hard to attach to it.

1

Honeymoon Objectivity

In 2014, a baby-faced, twenty-two-year-old entrepreneur named James Proud crowdfunded a sleep-tracking device that promised to automatically monitor sleep patterns, provide a numerical score, and make recommendations for sleep behavior. That such functions were already available did not escape Proud. Beddit, a sleep sensor that we will revisit in chapter 4, had been crowdfunded a year before and already released to its backers. In response, Proud chose to emphasize his device's "simple, uncomplicated and useful" qualities; designed as a slick, minimalist off-white orb, it would merge invisibly into the everyday flow of attention and reflection. "We believe technology needs to disappear," said Proud; "everything in [our device] is just designed to fade away."[1] It would carry an equally simple and no-brainer name: Sense.

In 2017, James Proud, now twenty-five, announced the end of Sense.[2] Panned by some tech reviewers as a "fundamentally useless" object[3] and a glorified alarm clock, the device never quite delivered the quiet transformation of everyday life that its creator aspired to. Fundamentally, it proved not very good at making sense of human sleep. Users reported that any deviation from the presumed sleep scenario—for instance, a pet snuggling up in bed—would throw the device off entirely. The chaos of everyday life rarely conformed to the expectations of the tracking machine, even as its selling point was that it would discover truths about us that we cannot perceive ourselves. As Proud's team wound down operations, users began to report that their Senses were losing functionality. The orbs went mute and deaf to the data around them, a small monument to the unfulfilled promises of new technologies.

Technologies of datafication reconfigure what counts as truth and who—or what—has the right to produce it, and not simply through the success of indisputably superior machines or even their mundane

and the institutions and machines that are meant to know it in their stead, between the practical capabilities of data technologies and the wider fantasies that give them legitimacy. In each one, we find troubling asymmetries in how different bodies are treated to different kinds of factmaking. If data expands the vistas of human action and judgment, it also obscures them, leaving human subjects to work ever harder to remain legible and legitimate to the machines whose judgment they cannot understand. Caught in an expanding and consolidating data market, we cannot simply seek more and better knowledge but must rethink the basic virtues and assumptions embedded in that very word. What kind of good does knowing do? Or, rather, what must our knowledge look like that it may do good? And who are we, with what kinds of capabilities and responsibilities, with what role to play in a data-driven society? As the truth of who we are and what is good for us is increasingly taken outside ourselves and human experience, the figure of the human subject—which, Foucault had warned, is a young and temporary thing[15]—is flickering uncertainly, unsure of the agency and moral responsibility we had worked so hard to attach to it.

1

Honeymoon Objectivity

In 2014, a baby-faced, twenty-two-year-old entrepreneur named James Proud crowdfunded a sleep-tracking device that promised to automatically monitor sleep patterns, provide a numerical score, and make recommendations for sleep behavior. That such functions were already available did not escape Proud. Beddit, a sleep sensor that we will revisit in chapter 4, had been crowdfunded a year before and already released to its backers. In response, Proud chose to emphasize his device's "simple, uncomplicated and useful" qualities; designed as a slick, minimalist off-white orb, it would merge invisibly into the everyday flow of attention and reflection. "We believe technology needs to disappear," said Proud; "everything in [our device] is just designed to fade away."[1] It would carry an equally simple and no-brainer name: Sense.

In 2017, James Proud, now twenty-five, announced the end of Sense.[2] Panned by some tech reviewers as a "fundamentally useless" object[3] and a glorified alarm clock, the device never quite delivered the quiet transformation of everyday life that its creator aspired to. Fundamentally, it proved not very good at making sense of human sleep. Users reported that any deviation from the presumed sleep scenario—for instance, a pet snuggling up in bed—would throw the device off entirely. The chaos of everyday life rarely conformed to the expectations of the tracking machine, even as its selling point was that it would discover truths about us that we cannot perceive ourselves. As Proud's team wound down operations, users began to report that their Senses were losing functionality. The orbs went mute and deaf to the data around them, a small monument to the unfulfilled promises of new technologies.

Technologies of datafication reconfigure what counts as truth and who—or what—has the right to produce it, and not simply through the success of indisputably superior machines or even their mundane

and ordinary operations in concrete practice. As relatively nascent technologies fast-tracked to the status of a global buzzword, the very idea of big data and smart machines—the spectacular keynotes and product launches, the anticipatory rhetoric, the science fiction, the projected future functions—operates as a social actor in its own right.

Proud's story, after all, is a common one. The popularization of self-surveillance technologies followed decades of anticipation (and disappointment) about a future that was always advertised as just about to arrive—a "proximate future."[4] A computing magazine included "smart appliances" in a 2007 piece about the "biggest technology flops," deriding the "bubble" around smart appliances back at the turn of the century. "The bubble burst, and we haven't heard much about intelligent appliances since," the article said.[5] That very year, the Quantified Self (QS) community would emerge in Silicon Valley; by 2011, the Internet of Things was back on the forefront of the imagined future, featuring in the tech advisory firm Gartner's influential "Hype Cycle for Emerging Technologies" report for the year.[6] Yet this return to the spotlight of the imminent future was not necessarily built on clear and proven cases of better knowledge. The rapid growth of the self-surveillance industry provoked public skepticism, academic research alleging negligible or backfiring effects,[7] and even lawsuits challenging the basic accuracy of popular measuring devices (namely, *Brian H. Robb v. Fitbit Inc., et al.* 2016). The broader industry of smart machines was no better off; one internet-connected juice maker cost $400, but its proprietary juice packs turned out to be just as squeezable by hand. A smart lock automatically updated over wireless connections and then locked users out of their homes; smart salt shakers promised voice-activated controls but were unable to grind salt. The proximate future was cobbled out of Eric Kluitenberg has called imaginary media:[8] prototypes depicting impossible realities, products sold on the basis of never-quite-actualized functions, artists' sketches, and bullish press conferences. Even as they malfunction and disappoint, they help drag impossible functions and nonexistent relations into the realm of the sayable and thinkable. Consumers are asked to buy into the latest gadget in anticipation of its future ubiquity, to install software for its future functions, and to celebrate prototypes for their glimpse of what, surely, must be just around the corner.

Technologies of datafication seize the authority to speak truth not by achieving the improbable hype gathered around it but by leveraging those lofty goals to mobilize the public, siphon funding, and converge collective imagination. Technology thus operates not merely as tools and functions but also as a panoply of fantasies—about machines that know us better than we know ourselves, about predicting the future with pinpoint accuracy, and about posthuman cyborgs and Big Brothers. To say "fantasy" does nothing to undermine the unique importance of material facts (as if fantasies could be sustained, or even generated in the first place, without the affordances of concrete things!). But it does mean tracing the ways in which data-driven surveillance seized its claim to knowledge by mobilizing projections and estimations about technology and the future world that will necessitate those technologies. While tracking devices such as Proud's were crafting an optimistic technofuture animated by consumerism, tugging on the broader imaginary of posthuman augmentation, state surveillance systems were warning of a future that must not happen, predictions of crime and terror that must be snuffed out through strategies of "zero tolerance." Across both cases, fantasy takes half a step outside present reality not to escape from it but to all the more effectively guide it.[9] Žižek once observed of the dystopian science-fiction film *Children of Men*[10] that

> Hegel in his aesthetics says that a good portrayal looks more like the person who is portrayed than the person itself. A good portrayal is more you than you are yourself. And I think this is what the film does with our reality.

The market projections, promotional rhetoric, bullish claims, and dystopian warnings surrounding datafication today are precisely the little doses of fiction used to make sense of these technologies and the knowledge they promise. Such beliefs are not reducible to "intellectual mistakes" by naïve or ignorant subjects. This (mis)recognition of what technology does and could do, the benefit of the doubt and the doubtful benefits, is so often a crucial part of getting technoscience off the ground.[11]

When New Technofantasies Are Old

Round about stood such as inspired terror, shouting: Here
comes the New, it's all new, salute the new, be new like us!
And those who heard, heard nothing but their shouts, but
those who saw, saw such as were not shouting.
—Bertolt Brecht, *Parade of the Old New*

Fantasy, in this collective, commercialized, politicized form, is never
truly free-form. Datafication often falls lockstep with familiar narra-
tives around machines and rationality, tapping into that modern drive
to order the world as a taxonomy of facts for a sense of legitimacy and
plausibility. As Lauren Berlant shows, these familiar anchors help stitch
together the contradictions and disappointments of technology, the gaps
between knowledge and uncertainty, into a sense that "the world 'add[s]
up to something,'" even when that belief is constantly displaced and
disappointed.[12] The paradox throughout this book is that technologies
of datafication rely so heavily on the imagined legacy of the Enlight-
enment, and its particular alliance of objectivity, human reason, and
technological progress, even as its deployment threatens to destabilize
the presumed link between information, human Reason and democratic
freedoms. Since its emergence over the mid-nineteenth century, the
thoroughly modern concept of technology has depicted a world whose
every aspect stands ready to be flattened, standardized, and turned into
problems that the ceaseless march of new inventions would render into
objectively optimal states.[13] The fabrications explored in this book lever-
age what we might call honeymoon objectivity: the recurring hope that
with *this* generation of technological marvels, we shall establish a uni-
versal grounding for our knowledge, a bedrock of certainty, a genuine
route to the raw objective layer of the world around us. By invoking this
long quest, tracking technologies are able to draw together their own
imperfections, uncooperative material conditions, incompatible and
otherwise resistant humans into a seductive vision of better knowledge.

The objectivity invoked by data-driven surveillance constitutes no
rigid dogma but a sedimented range of attitudes and affects embracing
a distinctly modern way of thinking and feeling about knowledge. As
extensively chronicled by Lorraine Daston and Peter Galison,[14] older

renditions, such as the scholastic *obiectivus/obiectiv*, generally involved
definitions starkly different from the modern. Even in Kant, objective
validity meant general "forms of sensibility" that prepare experience,
while the "subjective" referred to specific and concretely empirical sen-
sations. It is only during the nineteenth century that the now-familiar
juxtaposition emerges: a neutral, "aperspectival" objectivity as the privi-
leged instrument toward truth and scientific inquiry and biased, unreli-
able subjectivity as its nemesis.[15] By the later nineteenth century, Daston
and Galison identify a dominance of "mechanical objectivity": a regula-
tive ideal that called for the elimination of the human observer from the
process of data visualization. The critical impulse for these conceptual-
izations was, of course, the advent of photographic technology, which
provoked new theories and standards for what counts as visual truth
and who (or what) might be best equipped to produce it. Photography
– despite its own long history of manipulation and contested meanings –
thus spurred new linkages between automation and objectivity, produc-
ing the ideal where "machines [would be] paragons of certain human
virtues" precisely by ridding themselves of human subjectivity.[16]

The public presentation of data-driven surveillance leverages these
older ideals of objectivity, and the cultural capital it had accumulated
through traditions of scientific inquiry. In self-surveillance's effort to
map the microbiome or record every moment of sleep, we find a con-
ception of the body as an aggregation of correlations. Health, produc-
tivity, and happiness are broken down into a set of hidden but logical
relationships that machines might read and catalogue—the same kind of
correlations that might help predict the lone wolf terrorist, enabling an
orderly distribution of risk and suspicion across the population. In this
cultural imaginary, the world is an indefinite archive, and the machines
of tomorrow, if not of today, will be up to the task of cataloguing it.

All this is not to say that the Enlightenment ever bequeathed a sin-
gular doctrine about technology and reason or that different practices
of datafication share a totally coherent conception of a value such as ob-
jectivity. As Lorraine Daston herself noted, each historical rendition of
objectivity expresses not some immutable quality rooted in natural law
but a mélange of aspirational values that happen to occupy (or, at least,
contest) a normative position at the time.[17] Indeed, honeymoon objec-
tivity describes that recurring tendency to *claim to* make new advances

toward such an immutable quality, even as the kinds of data actually produced might diverge significantly from this vision. Technologies of datafication do not subscribe neatly to any single definition but cobble together different popular imaginations of what objectivity looks and feels like. The central proposition of mechanical objectivity—the preference of nonhuman processes over subjective ones for the reliable production of knowledge—is retained as a basic article of faith, but one that is routinely transgressed and compromised in practice. The messy and flexible ways in which the virtue of objectivity is "localized" onto self-surveillance cultures reflects, above all, how broad and pliable the word has become. Like culture, objectivity exhibits a certain "strategic ambiguity."[18] Its many possible permutations allow a wide variety of interpretations and attitudes to rally behind a common banner, where more fine-grained definitions might have splintered them. Thus, the fantasy endures to pass on its allure to another institution, another machine.

Pure Data

If the pursuit of objectivity, in all its strategic ambiguity, is the well-advertised benchmark of data-driven surveillance, an equally crucial question is: What kind of regime of knowledge, what kind of social order, is it meant to deliver? This book argues that the many articulations of data's benefits, capacities, its factmaking powers, revolve around a mythologization of data as pure and purifying. This pattern emerges not so much in efforts toward the technical definition of data, but in the public discourse, where the very question of what data *is*—or, rather, what can data *do*—again involves a messy plurality of ideas and dispositions. Data, fact, information, and knowledge are often conflated such that they are either seen to naturally follow on from each other, bolstering a sense of legitimacy.[19]

Three years after the first Snowden leaks, BBC4 released a documentary titled *The Joy of Data*. Its host, the mathematician Hannah Fry, boiled it down to a pyramid. From bottom to top, she explained, data is "anything that when analysed becomes information, which in turn is the raw material for knowledge—the only true path to wisdom."[20] Fry left unsaid what exactly knowledge and wisdom were, but the hierarchical relation was clear: the raw objective facts gathered through new

technologies would serve as the foundation for better knowledge. This DIKW pyramid—data, information, knowledge, wisdom, in ascending order—is a fixture in many computer science textbooks. Underpinning it is a world in which everything we can or need know is reducible to positivist facts, and by descending to this atomic layer, we will be able to recover objective data for any problem.[21]

In this articulation, data and knowledge are inseparable bedfellows. Data is the ubiquitous ingredient in the buoyant dreams of better knowledge, the object unto which the hopes and fears of technological and epistemic possibility are invested. In its most elementary form, it is described as raw data: data generated by the machine but yet to undergo "secondary" processes of statistical analysis, cleaning, visualization, aggregation, and so on. It is data fresh out of the sensor, with no artificial additives. In this telling, raw data is seemingly anterior to analysis, classification, and attribution of meaning. The valorization of raw data is intimately connected to self-surveillance's vision of empowerment through objective knowledge. In 2015, one QSer suggested raw data access as one of the three "freedoms of personal data rights":[22]

> Without raw data, we are captive to the "interface" to data that a data holder provides. Raw data is the "source code" underlying this experience. Access to raw data is fundamental to giving us the freedom to use our data in other ways.

Similar sentiments were expressed by a host of prominent commentators, including QS co-founder Gary Wolf.[23] The widespread implication of raw data's nonmediated nature translates into the fantasy of data as a purifying agent: a technology that will produce knowledge stripped clean of politics, of human bias, and of troublesome differences in opinion and establish the clear and rational path forward. Yet, as numerous scholars have pointed out, the very idea of raw data is an oxymoron.[24] Data only becomes data through the human design of each algorithm, relational database, and deep learning system—although there are important differences in how much detail is determined by manual design and judgment and how much is left up to machine learning.[25] Data is no thing-in-itself that exists prior to observation but something

to be "achieved"[26] through a concerted process of production that can never rid itself of human subjectivity and sensibility.

The dangerous consequence of this aspiration to purity is that the human, social, historical, and moral aspects of data's fabrications are invisibilized—allowing familiar kinds of speculation and prejudice to reenter by the back door. Consider the effort to predict and intercept terrorists before they can cause harm. Chapter 5 examines known cases of sting operations where certain individuals—predominantly young, Muslim, of Arab descent, male—are marked out by state surveillance apparatuses for fabrication. Driven by a moral and political injunction to "zero tolerance," in which even a single terrorist attack is an unacceptable failure of prediction and control in the wake of the September 11 attacks, counterterrorist operations do not simply wait for the data but actively work to produce the necessary proof. Thus, in the case of Sami Osmakac, FBI undercover agents supply the individual with money and the weapons and explosives to be purchased with that money, and coach him each step of the way until arrest can be justified. Meanwhile, the Snowden files reveal the surveillance programs themselves to be inevitably human. Analysts from the National Security Agency (NSA) speak of "analysis paralysis" and the struggle to handle supermassive volumes of data, while placeholder entities, such as "Mohamed Badguy" and "Mohammed Raghead," for database-search interfaces reflect the all-too-human, all-too-crude underside of sophisticated data-driven systems. Criticisms of datafication have often invoked labels such as data doubles and doppelgängers to warn against how individual self-expression is being replaced by alternative identities recomposed from data extraction.[27] Alongside such "copies," we also find a variegated ecosystem of speculative entities: the Osmakac that might have been, the Raghead in the database. Here, datafication provides no mathematical certainty but a range of possible outcomes and correlations to legitimize highly anticipatory forms of surveillance, judgment, and incarceration. The desire for epistemic purity, of knowledge stripped of uncertainty and human guesswork, ends up with concrete practices that draw perilously close to the imaginations of purifying the nation and the body politic. Suspected terrorists, brown or white, religious fanatic or ethnonationalist, end up subject to very different forms of datafication even as the technology promises a neutral illumination of truth.

The idea that raw data can access an orderly and calculable reality stripped of historical forces, social constructions, and political disputes translates into a converse fantasy: that the individual body will be purified of the elements that impair its health and productivity, and the body politic too will be secured as a transparent whole. Yet a central argument of this book is that such a fantasy of epistemic purity—of knowledge untainted, of complete certainty—itself carries political and moral biases. The belief in raw and untainted data begets not only an excessive reliance on algorithmic factmaking but also extends the older and deeper cultural desire for sorting the world into stable and discrete pieces. The recurring temptation: What if we could predict and eliminate every bad apple, every violent individual, and every criminal intent? What if we could maximize everything good about our bodies, expel all the toxins, cut off the bad friendships, and optimize every habit? Just as the pursuit of better knowledge through datafication entails a social shift in what counts as objectively true, the collective faith in the purity of data entails using the data to try to bypass important political and moral questions, to try to purify bodies through technological solutions.

The Groundless Ground

The mythologization of pure data puts into ironic relief the original Latin: *data*, meaning "that which is given." Today, (raw, big) data's privileged position in objective inquiry and knowledge production seeks to normalize into the woodwork, becoming "something we would *not* want to deconstruct."[28] It has been called the "cathedral of computation," or a faith in "computationalism": the fantasy that data simply *is* and shall provide a reassuring grounding for everything else that trouble us.[29] This faith has immediate and practical rewards. If datafication promises objectivity, impartiality, and predictivity, all these epistemic characteristics add up to a valuable sense of *stability*. On one hand, specific processes of data-driven analytics work within narrowly defined parameters where inputs may be standardized, modeled, and otherwise manipulated. In other words, the algorithm's truth claim itself relies on a set of grounding assumptions about the world out there and its methodological relation to data—assumptions that it agrees not to question to get the job done. At the level of data as a broader, popular imaginary,

telling here is the enduring popularity of a rather naïve extrapolation of Shannon's law of information: the idea that we can progressively eliminate uncertainty in all situations through the addition of information (which are themselves certain), each of which would reduce uncertainty in varying amounts. Both as a technical procedure and as a social imaginary, datafication thus consists not simply of truth claims but also the normalization of a new kind of grounding for knowledge claims.

This grounding, this social basis of felt certainty, was precisely the subject of Wittgenstein's final, incomplete work. In it, he asks, "What is entailed in the simple phrase, 'I know'?" There is a curious masking function: the act of saying "I know this is a tree," for instance, does not establish any comprehensive or objectively certain proof that I really do know. Yet we trust such claims on a regular basis, tacitly agreeing not to question them too far; after all, only philosophers bother to hold regular debates revisiting whether trees really exist. Our knowledge claims provide no indisputable foundation. The very act of saying "I know" seeks to "exempt certain propositions from doubt," to agree to not to look too closely.[30] This infrastructure of common sense is what Wittgenstein provisionally labeled world-picture, *Weltbild*: models that allow us to cope with the world, to make certainty and judgment possible.[31]

Yet herein lies an unresolvable paradox at the heart of claims to better knowledge: the groundlessness of the ground itself, or, the ways in which the demarcation of what "counts" as good knowledge is ultimately arbitrary. Wittgenstein comments that "at the foundation of well-founded belief lies belief that is not founded"[32]—precisely because to claim "I know" is an act that removes its contents from the game of proof and justification. This arbitrariness is well exposed by young children not yet versed in the unspoken boundaries of the language game: "But how do you know it is a tree?" "Well, it has branches, a trunk, some leaves." "But how do you know those are branches?" "Well, if you look at an encyclopedia—" "But how do *they* know those are branches?" and so on until the frustrated adult snaps: "We just know, okay?" We might reasonably dispute whether such grounding is truly ground*less* or simply deferred and bracketed in sufficiently complex ways that it can be presumed in ordinary contexts. For our purposes, the two options have the same consequence. For ordinary subjects, navigating their everyday life, pressed to judge and form opinions about things increasingly be-

yond their phenomenological horizon (such as a vast and secret government surveillance system or bodily physiological processes beyond the human sensorium), there is a practical need to make or accept knowledge claims, to not question their indefinite regress as the child does. Hence, Wittgenstein inverts the typical model of proof to say that when I say I know this is a tree, "it is anchored in all my *questions and answers*, so anchored that I cannot touch it."[33] To question that it is a tree, or to question how I can know any such thing, shakes too much of the edifice built above it that it typically becomes *unreasonable* to question it.

It is this ground that is being reconfigured when counterterrorism efforts blur lines with sting operations or when self-surveillance technologies are promoted as superseding human memory, cognition, and other "natural" means of datafying the natural world and their own body.[34] The following chapters examine how specific and often-imperfect techniques for prediction and analysis become valorized as objectively superior knowledge. Meanwhile, a growing set of assumptions—about the nature of data, the value of human thought and machinic calculation, the knowability of the world out there and the human body as an information machine—become "set apart" and invisibilized, melding into the background of everyday experience and of public discourse on data-driven knowledge. Across both state and self-surveillance, the material objects of datafication constantly seek to sink into the background of lived experience—mirroring the disappearance of data as a social construction deep into the ground. The NSA's data collection occurs not at the embodied sites of personal communications but through undersea fiberoptic cables, restricted-access data centers deep in the Utah desert, or buildings hidden in plain sight as a brutalist New York skyscraper.[35] Self-surveillance devices, at first thrust into the spotlight as delightful novelties, are increasingly seeking to recede into the realm of habit and unnoticed ubiquity—where their influence on individuals no longer needs to be justified through active and spectacular use. Datafication, in short, seeks to become our groundless ground.

The groundless ground constantly encourages those who live on it to forget how contingent it is. Pointing to the most basic elements in scientific and mathematical reasoning, Ian Hacking speaks of "styles of reasoning": nothing even so complicated as a system of measurement or a law but something as elementary as, say, the "ordering of variety

by comparison and taxonomy."[36] Like the epistemic qualities Foucault charted in *The Order of Things*, these basic tendencies rarely come up for debate even as theories and ideologies are toppled. They change far more slowly and so provide stable grounding that allow us to perceive a fact as fact in the first place.[37] In the data-driven society, such styles of reasoning govern how we relate a number (an algorithmically generated expression of reality) to the body's sensory experience, to conscious human testimony, and to other sources of truth. It governs how bodies are turned into facts: what kinds of bodies become eligible for what kind of datafication and how different bodies are treated to different kinds of factmaking processes. To identify the groundless ground as ultimately arbitrary and conventional is not to say that they are therefore illegitimate; such fabrication is, once again, a normal part of the social existence of things.[38] What it does mean is that data's claim to better knowledge is not a given, and neither are the forms of factmaking they bequeath on society. There are important political and moral choices to be made around what kinds of authorities should serve as the groundless ground and what kinds of data, machines, and predictions should count as looking and feeling like truth.

The Data Market

The epistemic fantasies of datafication matter—not when or if they deliver on all their promises but in the present, where the mobilization of collective belief in those fantasies transform what counts as truth and certainty. The patterns and tendencies specific to contemporary state and self-surveillance stem from two important tendencies in big data analytics: indifference and recombinability. Big data analytics are predicated on the ceaseless production of data indifferent to its specific nature and without a rigid presumption of its utility—because this data will always remain open further exploratory analyses, recombining different datasets and analytical methods to discover unforeseen correlations.[39] This is indifference to causality in favor of correlation; indifference to "intelligence," in the sense that the data is collected without the prior establishment of an interpretive context; and, as subsequent chapters show, indifference to the human experience of the world and that context of everyday living. To be sure, indifference does not mean neutrality.

Even as many aspects of the analytical process become automated and left up to learning machines, the design of those learning processes and the initial identification of the kind of data to be gathered renders it an "interested," if not deliberately biased, process.[40]

One such driving interest is precisely the manufacture of usable, justifiable certainty. Algorithms, as Louise Amoore puts it, "allo[w] the indeterminacies of data to become a means of learning and making decisions."[41] Messy data, extracted from lived experience and social reality and reordered into machine-readable form and modeled into a comprehensible pattern, are leveraged to produce truth claims that are not simply true or false but are carefully packaged expressions of probability that harbor uncertainty by definition. These are deployed and sold as freely transportable systems for generating "insights" across different social problems. To begin with, technologies and products are often crafted for fairly specific purposes. But that very act of measuring often involves recombining whatever data that can be conveniently acquired until a useful correlation (i.e., a profitable payload) is discovered, and it is also common that such data collection later leads to new and formerly unimagined kinds of predictions. Thus, the sex-tracking app Spreadsheets measures "thrusts per minute," a largely pointless value for any human assessment of sexual intercourse but one that the movement sensors on a typical smartphone are well equipped to provide. Such sensors, originally implemented for distinct features (such as the use of accelerometer and gyroscope data to allow portrait/landscape orientations on smartphone screens), create new affordances for the business of tracking. Big data analytics often has "no clearly defined endpoints or values,"[42] precisely because its profitability hinges on the expectation that any given algorithm, any process of datafication, might potentially be exported as a standard procedure for an indefinite range of activities (and thus business opportunities).

State and self-surveillance, despite their many local differences, thus participate in a wider, cross-contextual data market. The seemingly technical tendencies of indifference and recombination work to encourage a particular set of political and economic realities. The optimism that any and every process can be improved through datafication constitutes a voracious impulse that reveals big data's fundamental affinity with capitalism's search for continual growth.[43] The larger the userbase, the more

data to be extracted, which not only refines the primary analytics but also increases the possibility of recombining that data for new uses (or for selling them on to third-party buyers). Thus PredPol, the prominent predictive analytics system for law enforcement, borrows from existing earthquake modeling techniques.[44] Palantir, a private data analytics company, was born out of funding from In-Q-Tel, the venture capital arm of the CIA, and then sold its products back to intelligence agencies.[45] It has subsequently begun to reach out to corporate clients, such as American Express and JPMorgan Chase, demonstrating the ease with which antiterrorist technologies and antiterrorist funding can be leveraged for civilian surveillance.[46] Fitbit, one of the most popular tracking devices during the mid-2010s, is piloting partnerships with insurance companies,[47] and a significant minority of products have been reported to share data with third parties,[48] following exactly in the footsteps of social media platforms' journey to profitability.

The data market advances what has been called "surveillance capitalism": the work of making the world more compatible with data extraction for recombinant value generation.[49] This perspective situates what is promoted as a technological breakthrough in a longer historical cycle of capitalist "logics of accumulation,"[50] including the postwar military-industrial complex.[51] In effect, the data market constitutes an early twenty-first-century answer to capitalism's search for new sources of surplus value. Here, new technological solutions are presented as (1) a universal optimizer, which is hoped to short-circuit existing relations of production and maximize the ratio at which labor power is converted into surplus value, and (2) itself a commodity, which may be hyped up for a new round of consumerist excitement.[52] Surplus value is located not so much in the optimization of prices and goods sales but in the optimization of data extraction and refinement.[53] The "profit" at the end of this process is sometimes obviously commercial, as in targeted advertising and the direct selling of consumer goods. But the profits or uses of surveillance capitalism must also be counted in the biopolitical sense, wherein state securitization seeks to identify and manage the normal population or the individual consumer is enjoined to render themselves more attractive to algorithmic decision-making systems through techniques of self-optimization. The constant traffic and recombination of data thus entail an ever-wider range of situations in which data may

substitute or override the claims of physical bodies, conscious subjects, and lived experience.

In commercialized spheres, such as self-surveillance (and even in state surveillance, where the drive to datafy produces opportunities for lucrative government contracts for private firms), the logic of accumulation is the engine that animates datafication's promise of better knowledge. In this light, the ongoing demotion of human knowing in favor of machinic measurement and data-driven insight is not simply an intellectual argument but a variation of what David Harvey called "accumulation by dispossession": the seizure of assets to release at extremely low costs, producing new opportunities for profit that predictably benefit those with incumbent capital.[54] The more devalued human intelligence, the better for selling artificial intelligence. With datafication, the deep somatic internality of the self—my desires, my intentions, my beliefs—are opened up for revaluation on terms distinctly favorable to new products and systems of datafication. Exhorting the virtues of self-surveillance requires downgrading the reliability of human memory and cognition, such that the smart machines—and the new industries of hardware sales as well as the subsequent recombination of that data—is seen as necessary to true self-knowledge.

These trends extend long-standing tendencies in the history of surveillance, both digital and otherwise. After all, Foucauldian discipline was never about the sovereign execution of coercive power through surveillance; it was itself a highly distributed and participatory practice pegged to the promises of knowledge and productivity. To be sure, embedded in the very word *surveillance*—composed from the French *sur* (above) and *veiller* (to watch)—is a specific relation: domination from "above" through optics. But alongside that straightforward image of Big Brother is a history of surveillance as a technique for producing truth, affixing subjects to the identities and roles prescribed by that truth, and, ultimately, *disciplining* subjects into general dispositions and ways of seeing. To ponder whether we are "panoptic" or "synoptic" or "post-panoptic" is to miss the broader continuity of that liberal principle in which subjects participate in their own surveillance through the internalization of a certain way of seeing.[55] The lesson shared across the panopticon, the ominous screens of *1984*, the highly visible CCTV installations in London's streets, is that what really matters is not (only)

the active relation of a watching subject and the watched one but the generalization of the condition of being *under* surveillance—a condition that corrals the human body and all it does into a standing-reserve of evidentiary material for interpretation, recombination, and classification.[56] From this vantage point, what is fundamental to surveillance is not the image of an Orwellian coercive control but a set of processes by which my truth becomes defined by those other than myself through a systematic and standardizing mode of organization. Surveillance, in this sense, is inseparable from the history of large-scale communication technologies and often develops in lockstep with the reach of the latter.

This book asks what kinds of politics, what kind of subjectivity, becomes afforded through the normalization of these technological fantasies around objectivity and purity and through the cross-contextual expansion of the data market. In the data-driven society, "what counts as knowledge" so often ends up a question of what counts as *my* body, *my* truth, *my* eligibility for social services, *my* chances of being targeted for surveillance, *my* chances at a job . . . Even as the idea of big data bloomed into a ubiquitous buzzword, its ambiguous consequences continued to break out in accidents and scandals. Some were told through the popular annals of outrageous stories: the man who was fired by algorithms,[57] the African Americans categorized as gorillas by Google Images.[58] Other controversies were more wide-ranging and enduring, such as the Snowden affair itself. It has been described as the "data wars": the growing social conflict over how people's algorithmic identities are determined and by whom.[59] Like the culture wars, what is at stake is the distribution of labels and associations by which we can identify, sort, and make judgments on individuals.

The trouble is that even as big data and smart machines invoke the thoroughly modern and Enlightenment imagery of technological progress and societal reform, this generalization of indifferent and recombinant factmaking often serves to retrench politics and economics as usual. The mix of naïve liberal individualism and technocracy that fuels the visions of machine-optimized futures provides no fresh political vision for the distribution of resources or the organization of collectives. There is only the conceit that with new technologies, we can finally achieve a fully automated luxury capitalism. Indeed, the very idea of "optimizing" reflects one of capitalism's essential assumptions: that there

is always another world beyond this one to plunder, that there is no end to expansion, and that we shall not run out of resources, of new conquests, new sources of value.[60] That capitalism, just like technology, just needs the next upgrade, the next invention, to really fulfill its pure vision of totally frictionless transactions and truly melt all that is solid into air. Dressed in the shining garb of technological novelty, datafication proves most of all the difficulty of proposing a coherent alternative to capitalism and the good liberal subject.[61] The push for datafication thus extends and depends on enduring fantasies around liberal values, even as its implementation often reprises old roadblocks and compromises. We now turn to one such impasse in the Snowden affair, where an unanswered question looms above all the debates around transparency and secrecy, surveillance and privacy: Can the public truly know for itself in the age of nonhuman technologies? If not, what kind of politics remains?