THE FUTURE OF DIGITAL SURVEILLANCE

WHY DIGITAL MONITORING WILL NEVER LOSE ITS APPEAL IN A WORLD OF ALGORITHM-DRIVEN AI

YONG JIN PARK

The Future of Digital Surveillance

The Future of Digital Surveillance

Why Digital Monitoring Will Never Lose Its Appeal in a World of Algorithm-Driven AI

Yong Jin Park

University of Michigan Press
Ann Arbor

For questions or permissions, please contact um.press.perms@umich.edu

Published in the United States of America by
the University of Michigan Press
Manufactured in the United States of America
Printed on acid-free paper

First published May 2021

A CIP catalog record for this book is available from the British Library.

ISBN 978-0-472-07484-6 (hardcover : alk. paper)
ISBN 978-0-472-05484-8 (paper : alk. paper)
ISBN 978-0-472-12882-2 (ebook)

For Mom

Contents

Digital materials related to this title can be found on the Fulcrum platform via the following citable URL https://doi.org/10.3998/mpub.10211441

Figures

Tables

Preface

Are humans hardwired for delicate decision-making in the management of public-private lives? Or are we helpless victims of surveillance as consuming digital media presuppose the loss of privacy? Exploring the oscillation between the tyranny of digital surveillance and the ideal of privacy, this book traces the origins of behavioral constraints and institutional impulses for personal data collection in digital technologies, such as the web, smartphones, and the artificial intelligence (AI) embedded in social network sites, search engines, mobile apps, and email. This book is a critical account of how and why surveillance is natural and privacy is not, and it argues, quite logically, that privacy will need to be socially constructed with a forceful collective will.

By dissecting the puzzles surrounding technologies, this book identifies how to bolster the conditional nature of human rationality in the dynamic interdependence between institutions and individuals. This book resists a technologically deterministic view—that is, digital technologies by nature do not cause surveillance, nor do they constrain individual privacy rights. Instead, the shaping of surveillance technologies is embedded in a complex set of individual social psychologies, institutional behaviors, and policy principles, with digital technologies outpacing our imagination. Stay tuned for the results. The future is likely to be from rosy, as digital surveillance will never lose its seductive appeal in the world of "normalized" AI, algorithms, robotics, and Web 3.0.

Acknowledgments

An irony: I disclose my personal emotions in a book describing privacy. Or maybe it is not so ironic after all, as we humans develop affection by sharing the personal side of our stories. We call those moments of opening our identities special, and through them we learn how to connect with colleagues, neighbors, friends—and love. Accordingly, my disclosure in this space is warranted, though with careful discretion.

This book was conceived when I spent a sabbatical semester at the Center for Information Technology Policy (CITP), Princeton, where I became fascinated by the resurgent interest on artificial intelligence (AI) and machine learning. Even before that experience, I had often pondered the mysteries of privacy and conducted numerous quantitative and qualitative investigations. The genealogy of privacy-surveillance scholarship runs deeper and wider than perhaps anyone can imagine. We often call it a *privacy* study if it focuses on individuals. It becomes a *surveillance* study if the emphasis is on institutional arrangements for power, control, and resistance. If we ask how various disciplines apply their lenses, we note that sociologists' concern may be the unequal distribution of capital associated with privacy skills, but political scientists may explore the power dynamics related to the privacy of one's identities. Psychologists, on the other hand, might take a closer look at individuals' cognitive calculus for trading off privacy. Even religious studies has its place in the privacy debate, as privacy emerges as a human issue when Adam and Eve realize the need to wear clothes—the first moment in which a line was drawn between the public and the private.

Today, distinguishing the private from the public has moved to the

web and AI. This book appears when we are on the threshold of a revolution in which algorithms, AI, the internet of things (IoT), and robotics will redefine our sense of privacy. I argue in this book that the wisdom of our collective response to the technological change will determine whether we will helplessly veer toward digital surveillance or reshape that force to protect privacy. A few years back I had a chance to visit the Facebook headquarters in Silicon Valley. The tech-savvy, young, and brilliant workforce impressed me. An upcoming clash between a West Coast code of technological ingenuity and an East Coast code of bureaucratic government regulation was foreseeable. But in what degree would the clash occur?

This book was inspired by numerous scholars whom I wish I could emulate. Genuine admiration goes to W. Russ Neuman, a political sociologist, whose encyclopedic knowledge and MIT Media Lab experiences taught me much. During the most difficult time of my life he guided me with all manner of advice. His book *The Future of the Mass Audience* is a lifetime source of knowledge, inspiration, and emulation, and this book is my first attempt at emulation. I am also indebted to the following scholars: Oscar Gandy, Eszter Hargittai, Jerry Kang, Wenhong Chen, Anabel Quan-Haase, Craig Scott, Scott Campbell, Rich Ling, Nojin Kwak, Marko Skoric, Phil Napoli, Dan Yanich, Jeong Nam Kim, Michael Yan, and Steve Jackson. Scholars and thinkers I have not had a chance to know personally but who have tremendously influenced my work include Nissenbaum, Acquisti, boyd, J. van Dijk, Couldry, Mosco, Dutton, Wellman, Rice, J. Katz, S. Jones, Turow, L. Humphreys, Helsper, Hampton, Livingstone, Gillespie, K. Crawford, Neff, Ananny, Sandvig, Metzger, Ellison, Mueller, T. Wu, Lyon, and the late Ithiel de Sola Pool. And I thank three anonymous reviewers and Dr. Elizabeth Demers at the University of Michigan Press for their critical readings of the drafts of this book.

This book is a technological forecast, looking back to look ahead at a special case of privacy. Efforts toward privacy will be snail-like in speed and modest in scope; the wave of surveillance will be unstoppable because surveillance is a natural consequence of our environment, as this book argues. Corporate players and institution will be decisive in determining the technological use and applications of data, and will resist change. Individual citizens-users, on the other hand, will remain puzzled and awkward in their responses, with limited knowledge of privacy-surveillance issues and little ability to deal with the force of surveillance. In the end, surveillance is natural, and privacy is not. This delicate puzzle is one I attempt to unlock,

permitting us to peek into the future of algorithms, AI, and the failure of the "laissez-faire" self-regulatory tradition of free enterprise. Understanding this delicacy will turn out to be the key to the puzzle of privacy.

A first book is special, like being in love or witnessing the birth of one's child. I would like to share this private moment of puzzle and wonder with Hyun Doo Cho, Seung Hyun Choi, Doo Hee Park, S Mo Jones-Jang, Hoon Lee, Sung Hee Joo, Grace Yang, Yu Won Oh, and Damhee Kim. I am grateful to Jae Eun Chung, Yoonmo Sang, and wonderful colleagues at Howard University. I value my friendship with Eunjoo Suh, who is to be credited with the birth of my family of Roland, Dainna, and Michelle Kyung Eun Son. I thank them dearly. Now that I reveal my gratitude, adoration, and personal attachments without their consent, may this book be a gift in acknowledgment of their love, support, and inspiration—a mysterious private feeling that often bears on that public topic called privacy.

PART I

Introduction

CHAPTER 1

Putting Individual and Economic Determinants in Perspective

> Everyone lives in a social encounter, involving him either in face-to-face or mediated contact with other participants. In each of these contacts, he tends to act out what is sometimes called a line—that is, a pattern of verbal and nonverbal acts by which he expresses his view of the situation and through this his evaluation of the participants, especially himself. Regardless of whether a person intends to take a line, he will find that he has done so in effect.
>
> —Erving Goffman, 1967

1.1. The Battle over Control of Digital Privacy

Surveillance is an inseparable part of human lives. Surveillance existed in primitive societies in a form of gossip, peeping, or eavesdropping that satisfied human curiosity. In the past it was a means to detect violations of social norms and rules in tightly knitted communities (Westin 1984, 1991). Church authorities in the medieval era also "watched" religious transgression and kept records of family units, namely for taxation (Schoeman 1984). The citizen's demand for greater privacy and freedom in general was a product of modern industrial societies where the weakening of church authority and the rise of secular state power led to increased individual freedom and choices, followed by the rise of anonymized social interactions in mass society (Inglehart and Wezel 2005). This book contends that the digital transformation of today, however, has reversed the trend and reintroduced the forces of surveillance of the earlier era even more powerfully. What will the future of digital surveillance look like? What will

be the future verdict on digital privacy control? And how do we determine a complex set of expectations and policy principles that will shape the future of digital technologies? These are decisive questions in the debates on privacy, and as it turns out, they are far from being resolved with the continuous explosion of digital surveillance technologies.

Consider the sharply divided positions revealed during the 2014 Munk Debate about privacy protection. The experts, debating the pros and cons of Edward Snowden's revelations about surveillance, articulated contradictory understandings of public interest and individual benefits concerning the use of personal data.

One position was defended by journalist and former human rights lawyer Glen Greenwald, assisted by Alexis Ohanian, an entrepreneur and founder of Reddit. Articulating the second position was former NSA-CIA director Michael Hayden, aided by prominent legal scholar Alan Dershowitz. Arguing that privacy is a fundamental individual right, Greenwald and Ohanian accused the National Security Agency (NSA) of a violation of individual rights. As they put it, the NSA was "eliminating privacy worldwide by collecting and storing all electronic communications that take place between all human beings on the planet. It is devoted to sweeping up every email, every telephone call, every Google search, every browsing activity, and every online transaction in which people engage." Hayden and Dershowitz, presenting themselves as the champions of public interest, shrugged off these accusations. They bluntly responded, "It depends on facts: what kind of purpose, what kind of danger, whether surveillance is a legitimate defense of freedom. . . . Surveillance properly conducted, properly limited, it is the balance. That is not government [we should worry about], but that is Google!"

The two positions are equally valid if each is understood from its own perspective. The irony is that the two positions share key assumptions in the debate over digital privacy rights:

- Privacy is a key ingredient in promoting individual liberty and democracy.
- Digital technologies pose unprecedented surveillance threats to the well-being of citizens.
- Commercial and public sectors increasingly blur the boundary between private and public.
- Government policy can play a constructive role in balancing the interests of surveillant and the one surveilled.

Even more fundamentally, both sides—whether for or against increased civil rights protection—built their arguments on an implicit agreement that surveillance is an unavoidable pitfall of digital lives. In doing so, they risked taking a McLuhanesque, deterministic position, by which digital technologies take a predetermined course of development into a surveillance machine (Czitrom 1982; McLuhan and Fiore 1967).

The question is not new. George Orwell in *1984* predicted the presence of a telescreen through which a totalitarian state could monitor every move by every citizen (see Neuman 1991, 2010). Orwell imagined the omnipresent telescreen would be directly wired into the Ministry of Truth, capturing and analyzing every word, every whisper, and every movement of citizens. This is a powerful metaphor that captures the notion of a ubiquitous surveillance system that collects, retains, and appropriates personal data. Orwell's *1984* portrays a fundamental battle between two forces—an individual citizen and a totalitarian system to control the information associated with one's political, social, economic, and cultural identities (Castells 1997, 2002; DiMaggio et al. 2001; Neuman 1991). What Orwell identified is the perennial tension between the surveillant and the surveilled; the subject with a gaze and the object under the gaze; those with power and those without power. It follows that analyzing the battle over the control of privacy and surveillance requires we have independent measures of (1) the agency of an ordinary person in resistance and (2) the system that conditions the individual's power. Accordingly, our recommendations for policy solutions must not get locked into rhetorical shrewdness in defense of one or another political or ideological position. Instead, this book argues that systematic analyses of the status of privacy entail social scientific inquiries with empirical evidence on the two components just identified.

1.1.1. From Web 1.0 to an Artificial Intelligence-Saturated World

The starting point of my analysis is the dialectical battle between (1) corporate institutions and (2) ordinary persons over the control of personal information. The analysis in this book will span the period from the mid-1990s to the year 2018; the discussion will then turn to today's personal data-based artificial intelligence (AI), embodied in a commercial digital ecosystem of Google search, YouTube, Facebook, Amazon, and so on.

The status of digital privacy rights hit rock bottom with Edward Snowden in 2013, and the continuous explosion of digital technologies

since the mid-1990s exacerbated the deterioration of civil rights protection. The US Federal Trade Commission (FTC)'s $5 billon fine against Facebook in 2019, for instance, is testimony, not of regulatory enforcement, but of how far an advertising-based commercial platform can defy a government agency by violating an ad hoc regulatory solution. It is important to note that the internet in that year had been dominant for less than a decade, with more than 80 percent of adults having access to the broadband in the United States (Madden 2014); the transition into mobile platforms had begun to take shape as more and more people became smartphone dependent. Wearable technologies were harbingers of the era of personal communication, as we become ever more dependent upon AI-driven decisions about what to do, what to think, and, ultimately, who we are (Park and Skoric 2017). Thus, this book does not confine itself to a technological artifact, a single company, or the anomaly of a particular period. Instead, the focus is on a holistic trend in which the explosion of digital technologies, such as mobile devices, smartphones, the internet, and algorithm AI, has challenged our sense of privacy.

The focus of this book resonates with the concern raised by Warren and Brandeis in their seminal Harvard Law Review article "The Right to Privacy" (1890). The intrusive technology of the day was a camera with a zoom lens that could photograph a person from a distance without her awareness. With this technology, the violation of privacy was most likely to occur in a one-to-one social interaction between private citizens. Yet the present day's technologies are far more complex, multifold, and subtle. New media companies such as Google, Yahoo, Facebook, and Amazon play multifaceted roles in digital transactions related to all aspects of political, social, economic, and cultural identities. The NSA surveillance program, Edward Snowden revealed, swept data concerning all types of communicative transaction, blurring the distinction between governmental surveillance activities and those by commercial firms. In this context, the hero of privacy and the villain of surveillance are not easily distinguished. Amazon founder Jeff Bezos revolutionized online shopping with the customized one-stop cart. But Bezos is for many observers a villain whose microtargeting innovations encouraged economic discrimination based on profiling (Gandy 2012; Nissenbaum 2004, 2009; Turow 2005). Silicon Valley companies like Google, Facebook, and Yahoo are said to have resisted the NSA's demands for data-sweeping, yet these tech giants are frontrunners in establishing centralized digital databases with concentrated market power.

Consequently, we should move beyond a simple dichotomy of privacy versus surveillance and take a nuanced approach, appreciating the oscillation between the two poles—a sort of pendulum swing between privacy and surveillance—with multiple angles that allow us to untangle the complexity of this issue.

In this book I define privacy cautiously. Privacy is a floating concept concerning one's control of information, encapsulated in personal data (and its flow) that can be associated with one's identities and digital selves (Goffman 1967). As we are dealing with a changing and moving target, so are digital technologies evolving, from Web 1.0 to Web 2.0, mobile devices, smartphones, the Internet of Things (IoT), robotics, algorithm AI, and so on. Individuals' ability to control data against unwarranted surveillance is a critical dimension of civic power (Castells 2002). Consequently, shaping digital surveillance technologies, within a complex set of individual social-psychological and institutional behaviors and policy principles, is a primary focus of this book.

1.1.2. Notes on Methodologies

Three methodological disciplines guide this book. The first is based on the conventional social scientific inquiries that thrive on the individual-level analysis of surveys. The second is a political economy perspective that dissects the institutional-corporate impetus behind the surveillance of personal information in digitalized networks. The third is a new-media policy analysis, with a historical and critical sense of how free-market-driven self-regulation of privacy protection has evolved in the United States. The key to all these lies in a rigorous social scientific approach anchored in hard empirical evidence, while branching out into multiple subdisciplines.

This interdisciplinary perspective, in which the unit of analysis ranges from individuals to institutions and policy, will be evident in this book's analysis of algorithm-driven AI. Examining the cases of Facebook / Cambridge Analytica and Amazon's smart home device, this book analyzes the unprecedented power of AI in translating a stream of personal data into automated machine-based decisions that exploit a person's private moments. Numerous privacy-surveillance studies most rely on monism—that is, a methodological singularity confined to a single unit of analysis (Neuman 1991). This book moves beyond monism; it argues that to analyze the

complexity of privacy without speculating on the basis of one's ideological commitment, one must have a holistic picture encompassing multiple academic disciplines merged at a higher level of analysis. In taking this approach, we risk losing the favor of a particular academic discipline, but we aim to reap the benefits of several interdisciplinary angles in tackling analytical subtleties.

The readers of this book might find it tempting to read each chapter quickly to grasp topical issues related to privacy and surveillance. Such an approach would work, as each chapter stands on its own. However, the book is structured to move from problems to applications and then to solutions. Part II, starting in chapter 2, discusses fundamentals of individuals, institutions, and regulatory policies (problems), while in Part III those insights are applied to understanding AI and its challenge to privacy (applications), with Part IV pondering policy options and alternative futures (solutions). In advancing the thesis of this book as a collective argument, I find it impossible to stay with one type of analysis without losing sight of the complexities the issues present.

The conclusions and recommendations at the end of this book are not to be taken as ultimatums, but should be treated as tentative, challengeable, and subject to change. New digital platforms in their algorithmic surveillance power and AI are already outpacing our imagination (Kuhn 1962; Pool 1983a), so malleable policies are necessary if we are not to be impeded by outmoded ideas. What will remain constant is the contour of a power dynamic that is rooted in the battle between institutions and individuals, the dialectical tension between these two forces (Giddens 1983), as this book defines the privacy debate. These forces, no doubt, will continue to attract analyses, by scholars and researchers, of social-psychological and behavioral constraints, as well as institutional impulses.

This brings us back to the 2014 Munk Debate, where the venomous rhetoric from the two sides seemed to deepen, rather than bridge, the political chasm between them. When the two teams had finished the debate, people in the room seemed surprised by the tenacity of each camp in holding to its initial positions. Some readers of this book may also be surprised, as it contends that the future of digital surveillance requires that we not shout at each other across an ideological chasm. We are about to engage in an elusive debate about the shape of our digital lives in the future of Web 3.0, artificial intelligence, and an algorithm-saturated world. Although we leave behind digital footprints in the course of our multifaceted lives

with political, social, economic, and cultural identities, I start with the issue of privacy in offline social interaction. Understanding human privacy behavior in its extradigital forms will provide invaluable insights on the challenges digital technologies pose to individuals and institutions in the management of public-private boundaries.

1.2. Theorizing Erving Goffman's Hunch

There is a great variation in how social scientists have understood the concept of individual privacy in human behavior. This is a debate that goes beyond definitions in the lexicon. Humans are innately social beings (i.e., we are *Homo socius*), naturally defining our identities in relation to others and drawing the line between public and private lives. How the boundaries of our private and public beings are defined, managed, and controlled—in the multiplicity of social interactions—is a subject for profound debate. This is the crux of understanding the management of privacy—not just for the present, but also for the future.

Erving Goffman (1967) is one of the first social scientists to outline the premise of privacy-related human behavior. He argued that a human life is dramaturgical and, using the model of a theater, analyzed frontstage and backstage behavior. Goffman offered the following premises on human self-presentation, summarized in Table 1.1:

- A person establishes a line between the public and the private, presenting several "selves" on life's stage, both formally and informally.
- A person is free to construct private and public "selves" but is confined by the stage on which one plays in a particular situation.

The front stage is where a person formally and publicly performs in everyday social interactions. In the closed setting of the back stage, people can be truly themselves, behaving differently than when on the front stage, privately and informally.

Goffman's premise is key in advancing our understanding of personal privacy into the digital era, when our sense of private-public boundaries is established, redefined, and erased by algorithms, AI, the Internet of Things, robotics, and so on. That is to say, personal data can carry numer-

ous personal identities associated with different social, political, economic, and cultural interactions, extending our selves onto a stage newly conditioned by digital platforms and not easily assigned to Goffman's front stage or back stage. It follows that one should be mindful of how to stage the "self" in relation to others, as well as how to interact with technologies and surrounding environments, while being able to constantly adjust the revelations of "selves" (see Pinch 2010 for the idea of a technological "merry-go-round").

Note the duality in public-private boundary management: there are selves (humans, agents) and structures (stages, theaters). This duality resonates with the idea that even a mundane interaction requires a delicate balance between a system (e.g., a merry-go-round) and a self (e.g., a rider, who manages her motion, posture, or distance from other spectators), but these performances are made possible and limited by the physical configurations of the system (Pinch 2010). In other words, a person, while freely exercising the will to be public or private, critically depends upon the external conditions of the stage, such as settings, other actors, and situations and episodes (Burke 1989).

Plato's allegory of the cave in *The Republic* also foresaw a person enshrouded by privacy, and we may imagine the cave as a private home. The cave symbolizes the structural confinement within which a passive and isolated individual rarely engages with the outside world (Tönnies 1887). But today the cave of the private home is constantly encroached upon by digital media, the internet, the computer, the smartphone (Neuman 1991). Here the digital technologies are the window connecting the person to the public. The dependence on digital technologies—as they effectively connect people to the outside—accelerates the blurring of private-public boundaries. In this way the footprints of every citizen, which used to be private but are now digital, have become increasingly available for others to collect, store, and appropriate.

Goffman (1967) suggested that human beings are hardwired for delicate decision-making as they interact with others in the management

TABLE 1.1 Dialectics of Public and Private

Stage	Boundaries	Context	Actor
Front	Public	Formal	Open
Back	Private	Informal	Closed

of public and private lives. Under what conditions do people rationally manage their public and private selves in their own interest? When the Gutenberg printing press opened up the possibility of democracy in 15th-century Europe (Pool 1983), literacy was the fundamental requirement for diffusing democratic ideas. Neuman (1991) has highlighted the centrality of literacy in promoting the participatory orientation made possible by the printing press, a technological intrusion into the private knowledge of the Bible monopolized by the institutionalized priesthood. This is remarkably different from the deterministic view of the passive citizens who populate Orwell's *1984* (Neuman 1991; Neuman et al. 2011). But there is a deliberative human condition that helps realize human rationality in interaction with a new technology.

Pinch (2010) expanded Goffman's idea of public self-presentation into a place where we interact with technology, painting a picture of a person carefully managing selves—on a merry-go-round in a theme park, a fun, mundane, staged setting where the person becomes part of everyone's spectacle. We can understand how a person manages, in a moment of fleeting joy, to perform the delicate task of handling selves. She may be staging her face and expression while maneuvering around the carousel's horse. And yet she is still consciously aware of surroundings, other people, and most important, a self. Here the overarching point is clear: human agency remains, and it is how individual interactions with technology are shaped by institutions that will push the surveillance properties of technology in a positive or a pessimistic direction.

1.2.1. Privacy Studies in Interdisciplinary Overview

Over the years, great works from scholars in information science, communication, behavioral economics, law and policy, sociology, and other fields have integrated diverse perspectives in an effort to understand the public/private dichotomy.

Trained as a philosopher, Nissenbaum argued that an outdated notion of privacy belonging to the citizen, whose impregnable "home is his castle" (2009, 111), permeates legal policy debates. According to Nissenbaum (2004), the castle metaphor invokes a physical world, making privacy a matter of seclusion in a space away from the unwanted gaze of others. The key to personal privacy control, she argues, is "contextual integrity"; that

is, a person's decision to be private or public depends on deliberate consideration of the context. A person who divulges health-related personal data to an electronic medical service may be unaware of potential exploitation by insurance companies, and thus has not considered the context. This is a significant violation of privacy because the person may have perceived the context for personal data use very differently. What Nissenbaum (2009) rightly pointed out here is the conditional nature of privacy as one moves in and out of distinct social domains. Individuals must deal with constant contextual erasures of public-private boundaries, an indefiniteness exacerbated by digital technologies such as algorithm-driven AI.

Similarly, Marwick and boyd (2014) argued that social media technologies collapse multiple professional and personal situations into a single context. Digital skill sets are therefore paramount in effectively navigating the blurred boundaries of private and public life. Sociologist Gary Marx (1998) punctuated a similar point, highlighting behavioral strategies for countering digital surveillance, among them avoidance, masking, and refusal to allow the release of data. Marx called those strategies "surveillance neutralization moves"—suggesting the ability to neutralize the power of machine surveillance. In this regard, Gandy (2010, 2012) also explained how the surveillance powers of digital technologies led to societal disadvantages and discriminatory practices. Gandy's work reveals a profound concern about a person's power to resist digital participation in health, politics, and e-commerce, warning of surveillance technologies that sort individuals into algorithm-based statistics.

These works did not directly address Erving Goffman's notion of public-private management. However, to restate our premise: the public and the private are determined by two forces in counterbalance: (1) the individual agency of self-presentation, resistance, and struggles against (2) the structural constraints of surveillance technologies. One's agency in public-private management never functions alone, but always in interaction with, always bounded by, institutional structures that take shape in response to technological stimuli. It is useful to contrast the cited observers' understandings with the metaphor of Orwell's telescreen—the ne plus ultra of constrained choice in presenting digital selves, an inescapable institutional setup in which individuals are forced to navigate their self-presentation and privacy. Collectively, what makes the preceding scholarly works important is their understandings of the oscillation between passive and active roles that people play (Marx 1988), as actions are situated,

contextualized, and constrained in multiple social circumstances (Giddens 1983; Nissenbaum 2004).

Multidisciplines in perspective

Psychologists have taken different perspectives on privacy as a symptom of individual emotional needs (Altman et al. 1981; Pedersen 1999; Schwartz 1968). Scholars like Ittelson (1970) have construed a desired level of disclosure or territorial security that influences individuals' relational outcomes, such as intimacy, anonymity, or isolation. When people develop social relationships, one's emotional desire to disclose private boundaries is essential to human behavior and even ontological being. Likewise, social penetration theory, deriving from developmental psychology, dissects how face-to-face (and group-to-group) interactions typically evolve from an initial encounter to a mature relationship, as personal boundaries peel off gradually.

Scholars in interpersonal communications also investigate privacy as a relational dimension (Bazarova 2012; Petronio 1991, 2010; Tidwell and Walter 2002). For example, communication privacy management theory (Caughlin and Petronio 2004; Petronio 2012) has been useful in understanding how privacy contributes to the development of interpersonal contexts, with a focus on how to assert one's boundaries to others and regulate the release of private information at an appropriate level. From this perspective, the boundaries shared between two people are a key to maintaining intimate interpersonal relationships. Accordingly, these studies focus on how to negotiate disclosures, when to allow private access to outsiders, and when to regulate how private information is shared.

Business, marketing, and information science literatures have different takes on privacy, often exploring consumers' willingness to reveal private information and the public attitude to partaking in personalization (Awad and Krishnan 2006; Earp and Baumer 2003). This is an understandable orientation in marketing-related disciplines, as the pursuit of effective e-commerce models at the organizational level became a paramount business interest with the advent of internet in the mid-1990s. On the opposite side of the spectrum, however, renowned political scientists (Etzioni 2007; Solove 2001) voiced their concern about government encroachment with the increased digitalization of surveillance power.[1] Closely aligned is legal scholarship (Kang 1998; Post 2001), which raises important questions

about the power struggle between the state and citizens, with the protection of privacy construed as the defense of citizenship from intrusive state power (Bennett 2011).

Table 1.2 summarizes the respective contexts and concerns of privacy debates within different disciplines. This is a simplified, thumbnail version of profoundly complex debates. Still, we can spot where these contributions leave room for future scholarship. Note the absence of a single unified notion of privacy across all disciplines. Different levels of analysis (namely, social, organizational, or individual) and units of analysis (social media companies, users, or policies) vary across disciplines, often leading to misunderstandings among scholars. This lack of conceptual uniformity reflects the intricacy of privacy-surveillance issues (see Anonymous 1998). This book is prompted by the complexity of the debate.

Moving on to the argument in this book, let me offer the following premises. First, the focus on interpersonal communication strategies and individual psychology has inadvertently left the institutional structure that constrains individual agency out of the debate over privacy. On the flip side, from the perspective of legal scholarship, an emphasis on the macro legal structure has left little place for individual resistance to power and agency. Some scholars (Barocas and Selbst 2016; Gandy 2012; Kang 1998), however, raised the problem of profiling and the potential social discrimination based on, for example, such as race or ethnicity. Renowned sociologists (DiMaggio et al. 2001; Fischer 1994; Giddens 1983; Neuman 1991, 2016; Pinch and Bijker 1984 for the social construction

TABLE 1.2 Disciplinary Differences in Dissecting Privacy and Surveillance

Discipline	Level of analysis	Interest at stake	Intrusive villain	Context
Interpersonal communication	Micro	Social interaction	Others	Face-to-face relationship
Psychology	Micro	Emotion, cognition, behavior	Self and others	Psychological relational needs
Marketing	Meso	Consumer behavior, willing to release data	Data misrepresentation, data security	Business, customization, advertising
Political science	Macro	Power, transparency	Government	Political protest, democracy
Sociology	Macro	Profiling, discrimination	Government, private companies	Social Change

of technologies) have also begun to question the dialectics between institutional constraint and individual agency, and the prospect of meaningful social change.

Julie Cohen's work is particularly illuminating, as her legal scholarship often defies disciplinary boundaries. She addresses how difficult it is to understand privacy self-management in terms of an individual's decision (2012a). In fact, the technological consumption of data, even when one consciously attempts to manage identities, results a chain of behavioral traces that go far beyond the immediate transaction. Similarly, Humphreys (2018), in her work on "the qualified self," examines the interaction between laypeople and the emergent structure of new media environment. She argues that sharing private details in digital platforms like social media and mobile apps must be understood not as a binary decision—either to be private or to go public—but as complex sets of personal moments of documenting one's self or selves in public events. Nevertheless, except in these works, very few systematic efforts have been made to integrate multiple disciplinary insights beyond the scope of a single academic discipline in silo. I provide simplified descriptions of the relevant studies in Table 1.2. I do not intend to claim that any of the respective perspectives are misplaced, only to point out that prior treatises may be insufficiently modeled with a unique disciplinary focus appropriate to the conceptual difficulties of privacy in our moment and in the future.

The foundational argument of this book is that privacy does not exist in a societal vacuum, since institutional conditions and individual agency interact to shape the direction of digital surveillance technologies. From this perspective, we can understand the subtle ways in which technologies are socially constructed and how such dynamics play out among corporate agents and individuals in determining the future of digital surveillance as embedded in communication technologies.

Figure 1.1 summarizes the organizational rationale aligned with the central thesis of this book. Readers may find it odd to keep returning to the dystopian image of isolated and powerless individuals facing digital technologies of surveillance, a mass society in which televised imagery stupefies gullible audiences (Ball-Rokeach and DeFleur 1976; Neuman 1991, 2016). But this is precisely the point of this book. As mass media are a potential threat to the well-being of citizens, I share a sense of a crisis and a fear of surveillance at the dawn of the AI-based technological revolution (see Williams 2005). But I base this book's conclusions on social scientific

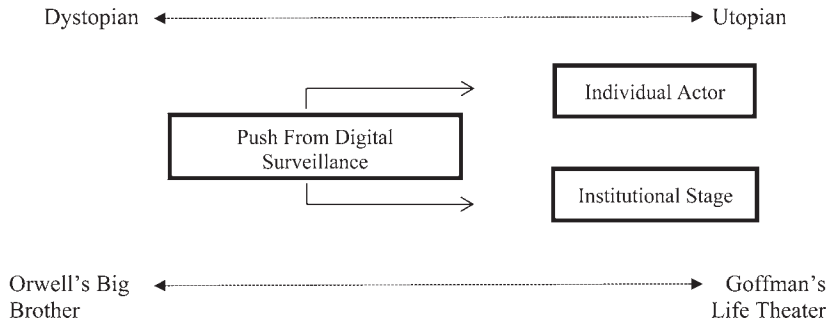

Figure 1.1. Interactive Forces in Tension. *Source*: Modified from Neuman 1991.

evidence and reasoning. The challenge is to understand to what extent the dystopian image of those surveilled remains valid in the digital era, and to discover if people can protect themselves and their data, rather than falling victim to digital surveillance. It will turn out that their selves remain deeply embedded in the institutional contexts into which individuals are "thrown" (Heidegger 1962). Yes, Google, Facebook, Amazon, and Yahoo have made our lives easier by enabling us to search for information instantly, by connecting us with old friends, family members, and colleagues, and by curating an endless barrage of news and information tailored to our personal interests. But there are personal data that we do not want to reveal and that if revealed may have unintended consequences. Some data are just meant to be private.

1.2.2. The Central Thesis and Organizational Logic of This Book

This book is a technological forecast, looking back to look ahead at a special case of privacy, and argues that moves toward privacy will be snail-like in speed and modest in scope, whereas the wave of surveillance will be unstoppable. This outcome is a natural consequence of our technological context if we let it develop as it is now developing at the institutional and individual levels. Yet we must avoid predicting the deterministic power of digital technologies with a blanket assertion of their inevitability. We should avoid the trap of either denying or accepting the inherent properties of digital surveillance (Neuman 1991; Neuman et al. 2011). The short

version of my forecast is: it will depend. No doubt, we do have a telescreen-like digital surveillance and passive citizens, as in *1984*. Further, as illustrated in chapter 2, the dramaturgical stage on which people play out their lives remains bounded by institutional imperatives that limit the potential of resistance, control, and effective privacy management. We might see a distant but potent possibility that surveillance threats can be reshaped by a few active individuals with the capacity to effectively navigate blurred private-public boundaries. But more fundamentally, the wisdom of our collective response to technological change will determine whether we veer helplessly toward comprehensive digital surveillance or actively reshape its force on behalf of our privacy.

Accordingly, I caution against blunt optimism that might neglect the powerful institutional forces that encroach upon personal privacy. But unqualified pessimism, which disregards the potential of individual power and resistance, is not valid either. Digital forces of surveillance do not determine the function of institutional systems and individual agency, but rather interact with them—hence the two-way interaction model in this book. The push by surveillance technologies should be understood as external to the interaction, and we start from this parsimonious assumption for the sake of analysis. However, dialectics between institutions and individuals contribute to the construction of technologies themselves. And it is critical to note that the forces of individuals and institutions in shaping digital technologies are not only interactive, but also countervailing or contradictory. The result is that George Orwell's dystopian *1984* meets with an idealized, utopian vision of life's theater offered by Erving Goffman, who expressed his confidence in people's innate abilities to manage private-public boundaries.

The organizational logic of this book is as follows: the first and most powerful force to consider is the economic power that ingrains collection, retention, and appropriation of personal data in the marketplace. This force, analyzed in chapter 2, responds to the push of digital technologies, but represents the political economics of selling and aggregating data in private and public sectors. The second interaction is the social psychology of privacy behavior, as individuals make use of digital technologies. This behavior, the subject of chapter 3, presents potential resistance to the intimidating push of digital surveillance technologies. This interactive dynamic, represented by the triangular diagram in Figure 1.1, is in progress as digital technologies continuously reshape our lives and thereby a

sense of privacy. In chapter 4, we visit this interaction from the perspective of regulatory policy and critically analyze US market-driven policy that conditions the interaction of the two primary forces.[2] The combination of these insights will provide a conceptual tool that can be used in chapter 5 for critical analysis of the function of AI. Chapter 6 suggests alternative evaluations of the puzzle of privacy, and chapter 7 concludes with predictions of the future.

The end of privacy as we have known it in the past is a certainty. This book appeals to its readers to appreciate the complexities involved in rescuing an individual right of privacy. Few citizens, even those in socially privileged positions, are well informed, and they are hardly resistant to the psychological enticements of comfort and the convenience offered by new developments in AI. To the extent that corporate institutions reduce people to binary inputs in the data ecosystem, they have already become constitutive parts of algorithmic AI systems that construct, regulate, and constrain individual behavior. Thus, without a regulatory blueprint for the fundamental change underway, we will lose the battle for privacy. We will see clearly—as this book makes it a plausible hypothesis—that the evil is embedded in a structural banality that promotes an inclination toward surveillance but against privacy. This process will be very hard to reverse.

The next chapter will show how the force of political economy responds to the push from surveillance technologies. We can predict a future in which the ultimate impacts of digital technologies will be determined by people who are enabled to resist surveillance and choose to control their privacy. Our point of departure is that telescreen-like, omnipotent surveillance is far from being our predetermined future, because there is no inherent reason technology must be constructed in that way.

PART II

Fundamentals of Privacy and Surveillance

CHAPTER 2

A Perspective on Institutions

> Always the eyes watching you and the voice enveloping you. Asleep or awake, working or eating, indoors or out of doors, in the bath or in bed—no escape. Nothing was your own except the few cubic centimetres inside your skull.
>
> —George Orwell, 1949

2.1. The Political Economy of Privacy and Personal Data Surveillance

George Orwell, in his novel *1984*, describes a desolate life under the ultimate surveillance machinery. The protagonist is Winston—an innocent citizen in a world entirely under observation. Big Brother, exercising the totalitarian power of surveillance, is present everywhere. Only thoughts, unwritten anywhere, seem to be beyond Big Brother's reach.

Some people today might argue that Big Brother has arrived and its name is Silicon Valley, with digital surveillance machine fueled by its powerful artificial intelligence. Others might go further, suggesting that even our thinking is no longer our own because the powerful algorithms in AI-based products customize the services and products offered to us and thereby dictate our behavior (see Gillespie 2014; Napoli 2015). Certainly, digital technology companies like Facebook, Google, Netflix, and Yahoo are whipping boys for privacy. For instance, Facebook and Google maintained a close relationship with the NSA up to 2013, when Edward Snowden blew the whistle on a massive surveillance program called PRISM. The companies have been charged with lying to the public about their relationship with the spy agency, gathering domestic intelligence for

the NSA and violating the Fourth Amendment, which prohibits any search without warrant. Nevertheless, in blaming the digital systems for the ills of surveillance, one must not confuse the characteristics of the symptom with its causes.

The thesis of this chapter is that there are structural and economic reasons for the deficiency of personal privacy. One of the central analytical purposes is to argue that the concerns about digital surveillance (the symptom) should be accompanied by a deeper understanding of its institutional impetus (the structural cause) if we wish to fix a broken system of privacy protection. The issue of functional impetus is critical to the understanding of a deep-rooted cause of institutional behavior, beyond the description of a symptom. The point is not to demonize the data practices of individual firms, but to diagnose the economic logic in the marketplace that induces data surveillance. From this perspective, the lack of privacy protection is incidental to the institutional structure that broadly frames the operation of digital systems. Accordingly, this chapter documents the existing industry structure, synergetic data practices, and the historical transformation of the past into a personal data economy, showing how these characteristics intensify personal data surveillance. A critical point of discussion is the extent to which corporate and institutional motivations, based on rationalized economic principles, remain at odds with normative societal goals such as privacy (see Hamilton 2000; Napoli 1999, 2001).

2.1.1. Intensification of Personal Data Economy

The digital personal data economy favors surveillance, not privacy. There are two main structural reasons for this. First, there is a macro-level trend toward market fragmentation, personalization, and niche marketing. Second, with the advent of the "big data" era, the incentive for digital audience measurement, such as real-time data collection, assessment, and appropriation, is accelerating (Hargittai and Marwick 2016; Neuman et al. 2014, 2016). These are the two intertwined institutional and financial practices fueling new forms of digital advertising.

Here the distinction between the digital technology industry and the media industry is important. It may be hard to pinpoint an exact moment in time, but at the very least, we can say that Silicon Valley and Hollywood have been converging since the early days of the internet in the mid-1990s.

Digital consumption via the Google search engine is a paramount example of this convergence, in which the expected roles of users, consumers, viewers, and citizens blur, with a collection of personal data enabling Google to paint a converged picture of any individual. An early equivalence may be found in the case of Netscape (purchased by AOL, Time Warner, and under Verizon as of 2019), the company credited with the implementation of browser cookies used to track users' online surfing.

Traditional media and the newer digital industry have evolved so as to converge with each another, along with the need for access to personal data across an increasing variety of digital platforms (Arsenault and Castells 2008). Fundamentally, the industry structure of the dual market, in which there are two distinct but interrelated markets for a single media product (that is, a viewer-user market in which digital consumption occurs and an advertising market in which digital advertising is sold), plays a critical role in the intensification of personal data surveillance (Napoli 1999, 2002; Neuman 1991; Webster and Ksiazek 2012).

Economic pressures toward fragmentation and personalization

The financial dynamics of traditional mass media center on the notion of the critical mass (Picard 2013; Vogel, H. 2014), which achieves efficiency by the sheer size of the audience. The larger the audience a program generates, the more profitable the media production will be. A "hit" in a television series tends to appeal not to the fringe, but to the broadest taste in the middle, as advertisers seek maximum exposure. At this point, there is a concentration curve in which a few hit TV programs recuperate the high sunk cost associated with the distribution and production of the rest of media products.

This drive toward homogenization and the concentration of hits is explained by the economic principle called Hotelling's centrism (Neuman 1991; Owen et al. 1978; Steiner 1952). Hotelling (1929) was primarily interested in the market behavior of two sellers. Imagine two competing stores in a small town. If the first seller, which is more popular, is located at the center of the town, the second seller will probably move closer to the center as well. If the first seller is successful in selling product A, then the second seller is likely to offer a similar product to attract more customers. The first seller will soon offer product B, which the second seller used to offer, to take back customers. The implication of Hotelling's centrism is

the inevitable predicament of product competition, that is, the excessive sameness of product offerings in the marketplace, where the interests of advertisers and sellers converge.

Figure 2.1, however, illustrates how digitalization has disrupted this traditional market dynamic. First, there have been drastic declines in the cost of transmission, storage, and processing of information, which has led to far greater number of sellers and buyers beyond geographically confined markets (Caironcross 1997; Couldry and Turow 2014; Neuman 1991, 2010; Shapiro and Varian 1998). Second, this has created the fragmentation of marketplace segments, where marketers can offer differentiated products with personalized tastes and pricing, not necessarily driven by the lowest common taste among the mass audience. Third, on the consumption side, the interactive properties of digital transformation have enabled individual consumers to self-select, search, and filter out media content, and to interact with the seller instantaneously for personalized services and products (Dutton and Peltu 1996). The traditional media model, which used to rely on big hits consumed by a large mass audience has been transformed into the economics of the broadband model, which thrives on the niche markets of a multitude of tastes and values.

For both advertisers and sellers, digital market fragmentation means understanding differentiated demands and being able to offer products and contents that will satisfy individual tastes. Data surveillance is the most efficient way to exploit market opportunities, because personal data provide access to consumption patterns and tastes, as well as psychographics, which can in turn be fed back into customized product service and pricing differentiation (Anderson 2004; Sunstein 2018; Turow 2005). Thus, personal data collection is not an intended violation of privacy, snooping into others' lives. Rather, data surveillance is the by-product of rational institutional behavior, purely based on the economics of business strategies that help marketers pursue lucrative values untapped in narrow market fringes.

The "long tail" thesis by Chris Anderson (2004) speaks to this fundamental logic, explaining why digital transformation is fueling data surveillance. The long-tail thesis posits that the 80/20 rule of Pareto's law, which states that 80 percent of the profit in the marketplace is represented by only 20 percent of what is offered, does not apply to digital environments. Anderson argued that the 80/20 curve is a product of inefficient distribution resulting from poor supply-demand matching in traditional mass

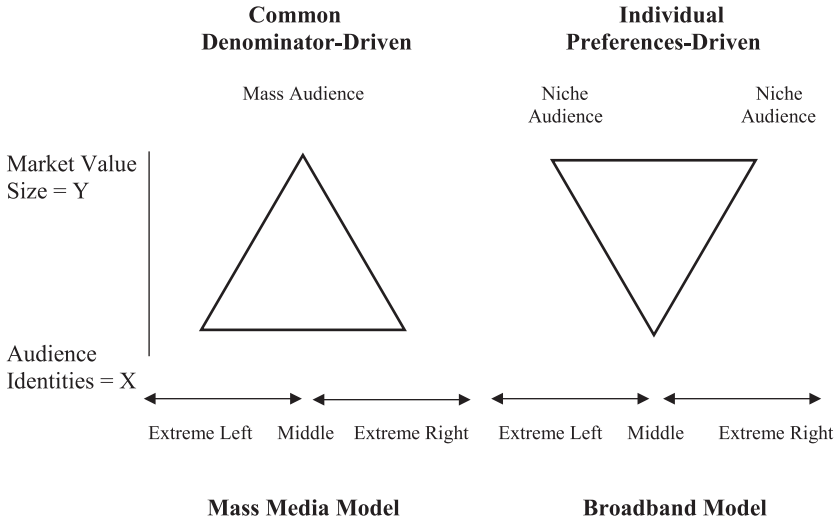

Figure 2.1. Digital Transformation of Business Model

markets—the inefficiency being exacerbated by the fact that most transactions are confined within one geographical region. Instead, a search engine like Google enables a nearly infinite number of small-scale transactions because it allows people to search for products with no geographical limit. More important, digitalization fixes any mismatch between buyers and sellers by helping individual users find specialized niches inexpensively. As we extend the long-tail logic, it is critical to see the perspectives of sellers. To them, personal data surveillance is nothing more than an effort to find a better match for their products by pinpointing ideal customers whose tastes suit their product offerings.

Note that criticism of the long-tail thesis derives from an empirical basis. That is, the combination of little pieces of "hits" (tail) do not add up to a big hit (the head). Despite an infinite number of blogs and multiple video platforms, the *New York Times* and the Super Bowl still draw the biggest audiences, thus offering the most expensive ad spots. Certainly, people's attention spans and powers of investigation are not broad enough in the vast digital universe to see beyond well-recognized brands such as the *New York Times*. Nor is it possible for a large number of individuals to be skilled enough to engage in diverse media activities and the consumption that numerous digital platforms may afford across fragmented market

segments (Park and Skoric 2017; Webster and Ksiazek 2012). Most scholars are therefore quick to dismiss the long-tail premise.

Still, the logic of the long tail remains valid to the extent that digital consumption patterns become increasingly differentiated, diversified, and personalized, with market values migrating from the center of a critical mass to the end points. Consider the case of Psy's "Gangnam Style." "Gangnam Style" generated the largest number of viewers in the history of YouTube up to the year 2017, thus propelling digital advertisers to flock to YouTube to place ads there for broad exposure. However, these viewers did not necessarily watch "Gangnam Style" at the same time or on the same platform. Instead, media consumption tends to be dispersed across tablets, smartphones, and laptops, resulting in another unprecedented competition among advertisers to migrate to the end points—that is, to track digital footprints across different platforms so that they can pinpoint individual lifestyles, localities, sociodemographic profiles, and associated psychographics (see Couldry and Turow 2014). Market value resides not necessarily in the total numbers of viewers, but in their particular characteristics, intensifying the pressure for data surveillance.

Economic pressures toward holistic digital audience measurement

One of the biggest population surveillances is the US census. The first US census took place in 1790 under the guidance of Secretary of State Thomas Jefferson, mainly for the purpose of taxation. As in most industrialized nations, the census in the United States is government-sanctioned surveillance, and because participation is mandatory, it has raised public concern about citizens' privacy (Singer et al. 1993). Scholars have questioned whether such a massive mechanism of collecting, processing, and retaining information regarding individual citizens exists for the sake of efficiency (Beniger 2009) or for the sake of bureaucratic social control (Gandy 1989; Kang 1998; Solove 2011).

In mass media, the best-known tool for population surveillance is the Nielsen rating system, and its precise roles have been a subject for debate (Napoli 1999, 2001; Streeter 1996). The fact that viewership, as manifested by audience rating, is an institutionalized form of currency in advertising markets is well established (Napoli 2012). Audience sizes result in the "Nielsen slope," according to which advertisers pay a premium for the highest-rated programs, whereas ads are sold at a discount rate for programs

with low ratings, such as public affairs programs and other unpopular content (Neuman 1991). Simply put, advertisers finance programs based on their audience rating—a measure of the size of the population who may be exposed to an advertisement attached to a particular program. In this way, the value of an audience is monitored and processed for different markets by Nielsen, which constitutes the centerpiece of media business.

Here one key lesson is that the economics of the mass-media industry evolve around the construction of audience valuation, that is, how a viewer or user population with certain sociodemographic characteristics will meet advertisers' needs. To put it differently, an institutionalized form of surveillance monitoring, such as Nielsen ratings, is the quintessential reflection of the media industry's structure, the economics of which rely heavily on audience valuation. Our conclusion draw from this evidence is that information on audiences, amid the digital transformation, will become an ever more valuable media market commodity (see Napoli 2012), intensifying the demand for an evidential basis on which behavioral patterns can be tracked with precision.

Extending this insight, a history of ratings in the United States shows how significant these ratings were to the improvement of efficiency and accuracy in audience surveillance measurement. The first measurement taken was a telephone survey in 1930, asking respondents to recall their own consumption behavior, followed by a telephone survey in the 1940s in which respondents were asked what program they were watching at the time the answered the phone. Starting in 1949, Arbitron controlled for potential sample bias by asking respondents to record their listening habits in a diary and mail it back. Nielsen's People Meter in the 1980s was an effort to move beyond the reliance on people's memories—a source of selective bias and recall error. The meter was an electronic data box that was hooked up to the television set to monitor, collect, and transmit viewing records to Nielsen—an enhanced surveillance machine that participants voluntarily signed up for in exchange for monetary compensation.

Our main points of this discussion are summarized in Figure 2.2, which highlights the contrast between mass-media and digital audience measurements. On one hand, there have been significant gaps between the audience that marketers want to reach, the measured audience that they think they are reaching, and the actual audience (Napoli 2012). These gaps represent threats to the validity of audience data, challenging the foundation of the traditional media industry and its information

regime. On the other hand, note the contrast with the broadband model, where the actual, the measured, and the imagined audiences are more likely to overlap because the entire user population conveniently falls into the sample frame. Theoretically, this means that there is no sampling error (SE) in audience measurement because of the direct observational data collection by which digital technologies provide a virtually 360-degree view of users (Danna and Gandy 2001). It is easy to see this if we think of the "big data" analytics that Facebook, Google, or Twitter can offer after having collected massive sets of personal data in petabytes of real time (Shapiro and Varian 1998).

Of course, it is simplistic to equate predictive social media algorithms to audience rating measurements in the broadcasting model (Bollier and Firestone 2010; see boyd and Crawford 2012). After all, social media users may not be representative of the general population, and the characteristics of a platform may exacerbate self-selection bias. Nevertheless, from the perspective of marketers, the large sets of detailed personal data and surveillance intelligence represent unprecedented advancement beyond the previous error-prone measurement of audiences, offering the capability of big data that precisely monitor audiences' activities in real time, with tools such as Like and Follow buttons and sharing of comments used to triangulate the data. And this is precisely the direction in which the AI-driven business model is evolving.

The Nielsen rating focuses on the size of audience, a one-dimensional measure of market composition that does not capture actual audience activities at the point of media exposure (Napoli 2011). The assumption is that exposure to advertising increases the probability that a consumer will be persuaded—whether the intended effect of the message is commercial, political, or a public service. The dubious effects of political and commercial advertising have been well documented and much debated (Schudson 2013), and the inefficiency of traditional audience measurement, on which broadcasting sectors still rely, is one of the factors to blame. As the costs of digital data collection and processing decrease, the institutional demand for personalized measurement, such as big-data analytics, is likely to increase so that advertisers can triangulate various data points and construct a holistic view of their customers. It is no surprise to see that digital platforms, which enable advertisers to reach actual users with greater confidence, are best positioned in fragmented marketplaces.

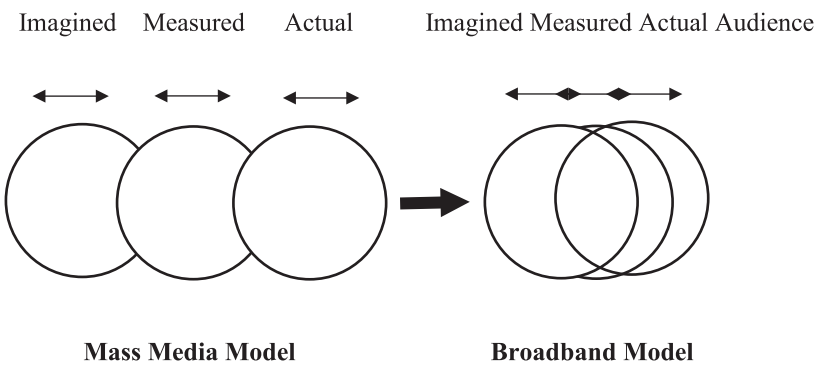

Figure 2.2. Digital Transformation of Audience Measurement. *Note*: Arrow denotes confidence range of audience-user measurement.

2.1.2. The Puzzling Economics of Privacy and Surveillance

When institutional motivations and causes are taken into consideration, the properties of digital transformation are fragmented and personalized; the costs of monitoring, collecting, and processing personal data drastically decrease; and with the advent of the big-data era, the parallel incentives for advancing digital audience measurement, such as real-time data collection, assessment, and appropriation, are accelerating. Theorizing the effects of digital surveillance technologies requires understanding the economic constraints under which commercial media platforms operate in audience marketplaces. The intensification of digital data surveillance is a product not just of technological change, but of structural adjustment by the industry, the economics of which depends upon the commodification of the audience (Hamilton 2000; Napoli 2001; Neuman 1991; Streeter 1996; Webster and Ksiazek 2012).

CEO Mark Zuckerberg once bragged about Facebook's surveillance capacity, that is, that Facebook could pinpoint individuals who would satisfy advertisers' needs better than any traditional media platform. Certainly, Wall Street investors who responded to Facebook's 2012 initial public offering appreciated its strategic advantages, as shares traded at $31.91 within the first week, the highest value in the history of the technology sector (Eichenwald 2013). Yet the fundamental concern remains similar to that with respect to the US census: the efficiency gain from surveillance

over citizens will involve the social cost of invasion of privacy (Beniger 2009; see Cohen 2012b; Gandy 1989; Solove 2001). As discussed earlier, the institutional effort to create more precise surveillance techniques is a logical step in repairing the invalid measurement of audiences. A point of contention is that economic efficiency, understood as the institutional effort to maximize output in the marketplace, does not conduce to protecting the privacy of audiences. From the point of the digital media industry, there is no institutionalized compensation for ensuring privacy, while surveillance has an immediate impact on the corporate bottom line. That payoff structure says a lot.

The conspicuous contrast between privacy and surveillance is illustrated in Figure 2.3, which displays a conceptual calculation for the benefit of a consumer-user-citizen, operationalized by a hypothetical interaction between privacy (here defined as an individual ability to control data) and surveillance (defined as data collection, retention, and appropriation) (see Park et al. 2018). Note the inverse relationship indicated by the black line, as a higher level of privacy is associated with a lower level of surveillance. On the other hand, the dotted line indicates a positive relationship between surveillance (again indicated by personal data collection, monitoring, and measurement) and personalization (indicated by digital media product, services, or content). Given that a higher level of personalization (see Couldry and Turow 2014; Neuman 1991) is required for increased surveillance, we see a somewhat paradoxical outcome—the individual benefit of privacy must amount to zero to achieve maximum personalization. In other words, the benefit of supply-demand transaction offered by personalization becomes available only on the condition that privacy remains low. On the demand side, this is an extremely inefficient marketplace because any transaction that involves personal data will never achieve a clear net gain (i.e., a gain in privacy means a loss in personalization, and the vice versa). On the supply side, the interaction is also inefficient, because there is no way for digital media producers, advertisers, and marketers to maximize the possible market output when a consumer-user-citizen desires both privacy and personalization. This is the puzzling economics of privacy.

The central point is that market economics may seem too crude to handle the delicate balance between privacy and surveillance. Here it might be foolish to predict that the greater the surveillance, the greater the personalization. But that is precisely the critical point. The fact is that data surveillance is justified on the basis of tangible benefits, such as personalized me-

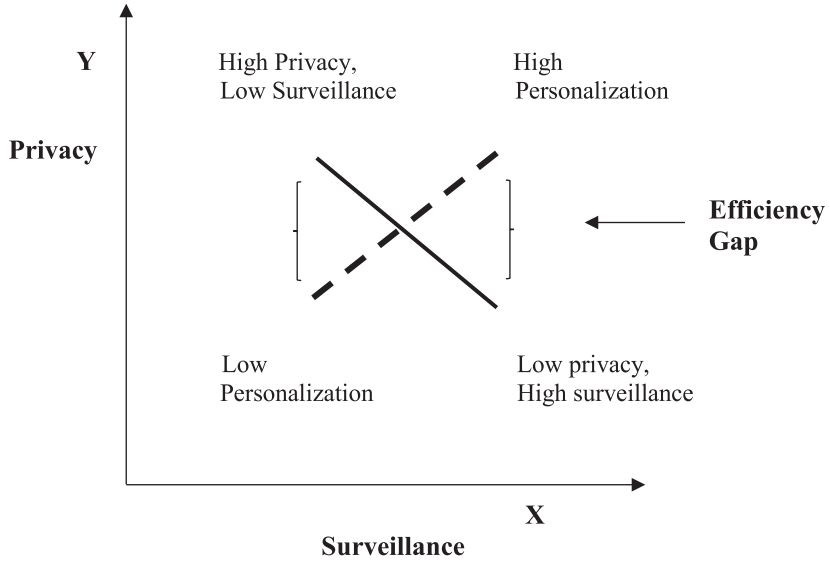

Figure 2.3. Efficiency Gap in Marketplace Performance. *Note*: Denotes the degree of interaction between privacy and surveillance denotes the degree of personalization. Here a cutoff point for low and high is only conceptual to illustrate the interaction.

dia services in commercial platforms or improved public safety produced by governmental surveillance. What I have illustrated in this section is that the market in which personalization and surveillance are highly correlated becomes highly inefficient when taking privacy into account. Earlier in this chapter, we discussed the alliance of the two pressures toward (1) fragmentation/personalization and (2) surveillance measurement, which accelerates with the digital transformation of the marketplace. Even when we assume that there is a perfect, clear-cut rationale for those institutional benefits, the marketplace alone does not seem to have a mechanism for ensuring privacy (Gandy 1989, 1998; Kang 1998; cf. Napoli 2001; Solove 2001).

2.2. Evidential Trends: The Locus of Privacy Protection in the Marketplace

In demonstrating the lack of a meaningful link between privacy and the marketplace, this section presents empirical evidence at different levels.

First, I examine the condition of personal privacy control in US. commercial websites and analyze the relationships between marketplace attributes and privacy protection, as shown in the interface design among the sampled websites. Second, beyond cross-sectional data, we can also observe longitudinal trends to detect whether certain privacy protections in digital marketplaces have improved over time. The potential mismatch between the provision of privacy protection (production side) and the public's concern with privacy (consumption side) is the third issue that demonstrates the function of the marketplace. Measurements were not precisely equivalent across the relevant studies used in this chapter, and sample sizes also vary, thus making consistent observation over time difficult. I caution readers against overinterpretation. Nevertheless, privacy and surveillance issues have led to systematic efforts to monitor the extent to which websites protect privacy. This section compiles evidence from published data sources (see Appendix A for the characteristics of sampled websites and methodologies).

The analytical focus of this section is the findings of published reports, starting with the landmark Federal Trade Commission Report of 1998, and then tracing patterns of privacy protection up to the year 2017. Adjustments are made to reanalyze data collected in 2008—the critical juncture of digital transformation when Web 2.0 companies, such as Facebook and Google, gained dominant market shares. This section asks how the digital marketplace appropriates privacy protection, and it presents a holistic picture of the two contrasting possibilities of the internet as (1) a commodity subject to data surveillance or (2) a platform for informed choice in managing personal privacy.

Privacy design for informed choice

Notice and Choice/Consent are the two categories of standards that the e-commerce industry and the Network Advertising Initiative, a consortium of internet advertising companies, adopted in the mid-1990s. The FTC originally recommended five standards (Notice, Choice, Security, Integrity, and Enforcement), already a reduced version of the FTC's own Fair Information Practice principles. Figure 2.4 shows how frequently the specified components of Notice and Choice appear among the sampled websites. The results show that privacy control features were insufficient in most websites. The overall provision was extremely low, with most items available in less than 10 percent of the sites.

Three items included in the privacy standard for adequate Notice are (1) availability of the policy on the home page, (2) correct labeling of the privacy policy, and (3) a different font color that distinguishes the policy statement from the background color. These were the only interface (notice) features present in more than 50 percent of the sites. All the other features of privacy control, however, were present in barely 10 percent, indicating that even the basic interface elements, such as placing privacy policy in the menu or having a different font color, were not commonly implemented in commercial websites. The sampled websites tended to provide more items that are part of the Choice privacy standard. For instance, 57 percent of the websites had an active email contact for privacy-related inquiries, and 26 percent of websites included a preference edit function. However, any meaningful privacy-related decision and action would not be plausible without such items as out-links to complain (for instance, to the FTC or industry advertising groups), an opt-out function, information on policies associated with disclosure to third parties, and so on.

Given that the financial costs of implementing these design elements are close to zero, the results of this survey indicate that the institutional willingness to protect personal privacy is limited. Digital industries do not apply these standards rigorously. The undersupply of privacy protection elements, as shown in the distribution of Notice and Choice features, reinforces the incongruence between privacy and marketplace.

Table 2.1 presents the results from bivariate correlations in Notice and Choice and the logistic regression model that analyzes the likelihood of high or low provision (coded as 1 or 0) of privacy protection elements. We can expect minimal or no effect of any characteristics of individual websites if privacy protection is not the function of marketplace. For analysis, the independent variables of market factor were divided into domain and site characteristics, and the result shows no significant relationship.

First, in terms of bivariate correlations, the more popular the websites were, the more Choice elements they provided. But there was no positive effect of website popularity on Notice. It was also found that the more recent websites were, the fewer Choice elements they provided, which suggests that newer sites did not perform better with respect to privacy. Those sites with more unique US users, indicative of market function directly under the current regulatory regime, tended to provide more *Notice* items, but the result did not hold true in logistic model when it took all other factors into account. Overall, the explanatory power of the logistic regression

as a full model remained very low (pseudo R2 = .11), with the only one variable (popularity ranking) displaying a positive effect. From the findings, it seems clear that marketplace factors made minimum contribution to the higher level of privacy protection. No clear benefit from the multivariate market factors was evident, as almost none of the domain and site indicators showed a significant effect. In sum, all of these considerations point to our conclusion that marketplace resources may not be readily translated into privacy protection and a tangible action that incorporates the demand for privacy.

The cross-sectional data by nature do not fully account for variations over time, and are not immune from yearly fluctuations. It is possible that nonmarket effects are the product of idiosyncratic factors in a particular year. In addition, market pressure for privacy protection may work gradually over time, concealing immediate effects. For this reason, it is crucial to make historical observations beyond cross-sectional data to make a better case for or against the market function related to privacy protection.

Privacy protection in over-time trend

Figure 2.5 shows the pattern of privacy protection, as indicated by three items, over a ten-year span. It is evident that even in the earlier days of the

TABLE 2.1 Bivariate Relationships between Market and Privacy Protection

	Notice *r*	Choice *r*	Notice + Choice B
Domain characteristics			
Ecommerce	.04	.01	.07
Online only	−.03	−.05	.21
Sensitive 1	.03	−.02	.89
Sensitive 2	−.03	.02	.83
Site characteristics			
Ranking	.03	.30 **	.59 **
Public	.02	.12	.34
New	−.07	−.25 **	−.09 *
US oriented	.14 *	−.04	.00
			Pseudo R^2 = 0.11
			X^2 = 18.92

* $p < .05$; ** $p < .01$

Note: Sensitive 1 indicates whether a site is targeted toward children, teenagers, or younger users (yes = 1, no = 0), with sensitive 2 being about whether a site deals with sensitive data (health or financial data) (yes = 1, no = 0).

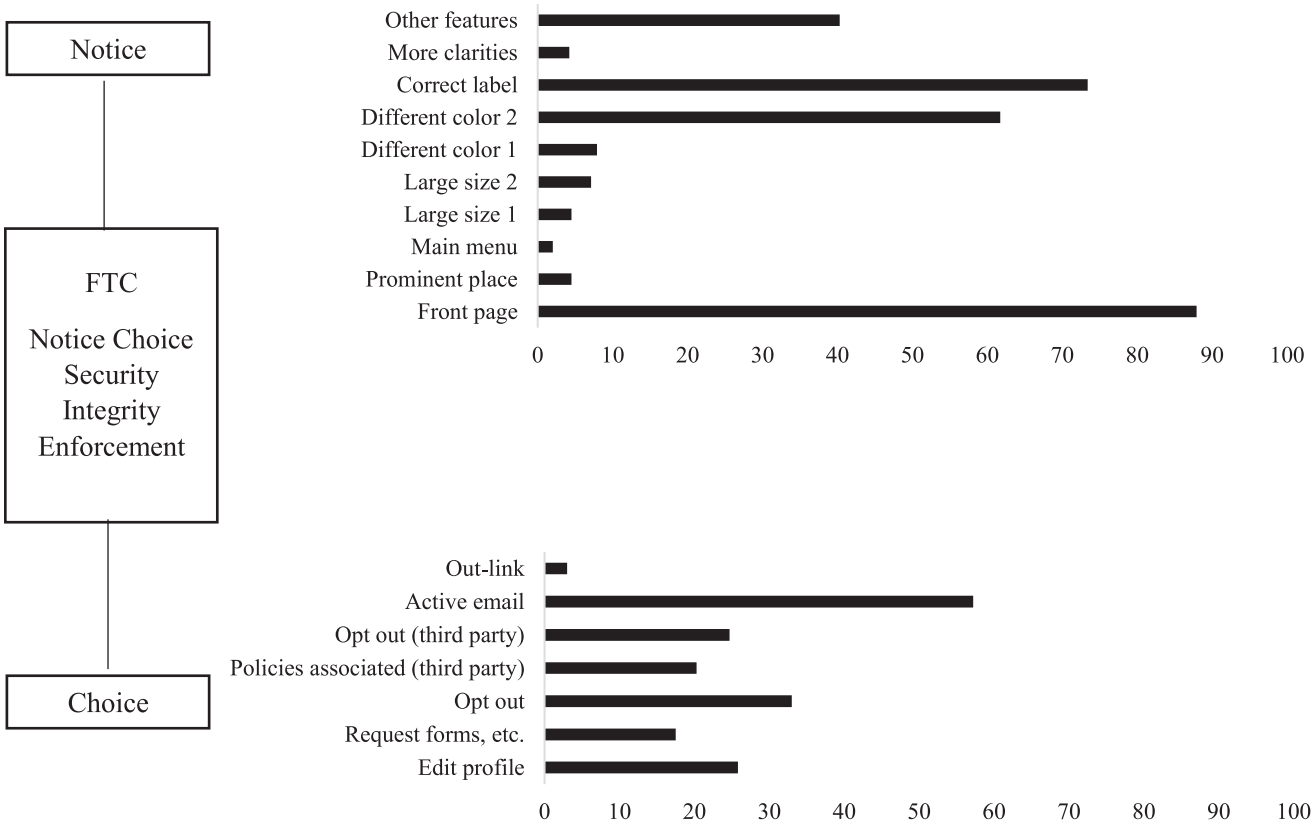

Figure 2.4. Marketplace Performance of Privacy Protection: Notice and Choice. *Note*: Percentage of the sampled websites shown.

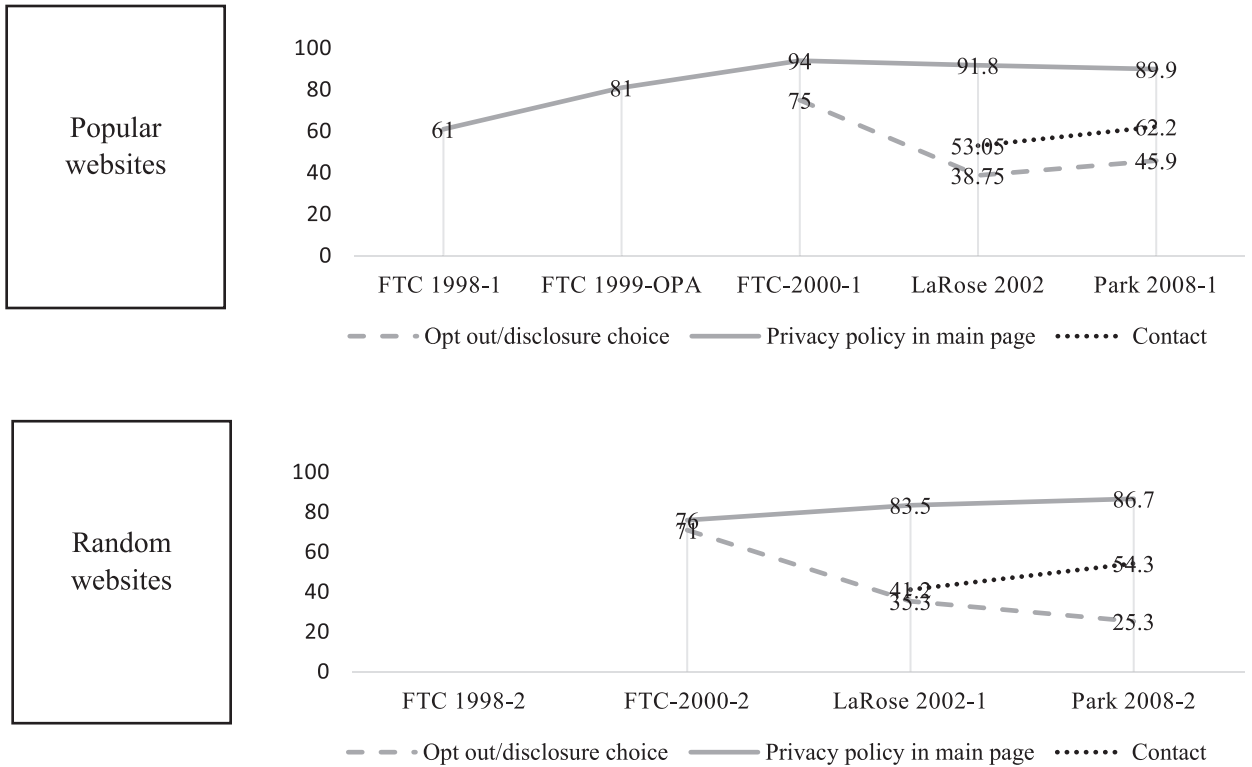

Figure 2.5. Marketplace Performance of Privacy Protection Over Time. *Note*: Percentage of the sampled websites in each study shown.

internet, a majority of popular (upper panel) and random (lower panel) websites provided privacy policies on their home pages. This is not be surprising given the low cost of placing a policy statement online. What is surprising is that not all sites have privacy policies on their main home pages. More important, the proportion of websites providing an opt-out choice decreased over time, while the provision of a contact option for questions on privacy only slightly increased. This indicates that there was no discernible over-time increase in the measures to protect privacy. Popular and random sites was remarkably similar in their provision of the three items, illustrating that a site's visibility and popularity may not result in the improvement of privacy protection over time.

Privacy demand (consumption) and privacy protection (production)

The central contention in this section is corroborated in two additional ways. First, it is important to track the level of public concern to detect whether institutional practices match any increase or decrease in concern over privacy. The logic is that privacy consumption (demand) and production (supply) should match if the marketplace is functioning beneficially to align consumers with suppliers. Second, it is important to detect privacy protection in recent digital technologies and platforms, as more recent observations will give us confidence that we are describing current practices. Our information should include not only website features but also integrated data services in different types of consumer devices.

Figure 2.6 shows the enduring incongruence between privacy protection (production side) and public concern (consumption side). The trajectory of concern over privacy is represented in the left panel, which shows that public anxiety doubled from 1998 to 2015, reaching a peak in 2014 after Edward Snowden broke the story about NSA surveillance. The right panel indicates the relative stability of privacy protection from 2002 to 2008. I averaged all three (popular and random) samples of the last two years from the available data set to estimate this trajectory. The ups and downs of public sentiment reflect political events, such as the 9/11 and the Snowden scandal in 2013. But protection of privacy does not match these fluctuations in public concern. That is to say, institutional practices in the marketplace do not seem to respond to increased public demand for privacy. What commercial websites responded to is not public concern over privacy, but a different market demand, namely, for better audience surveillance, as reviewed in section 2.1.

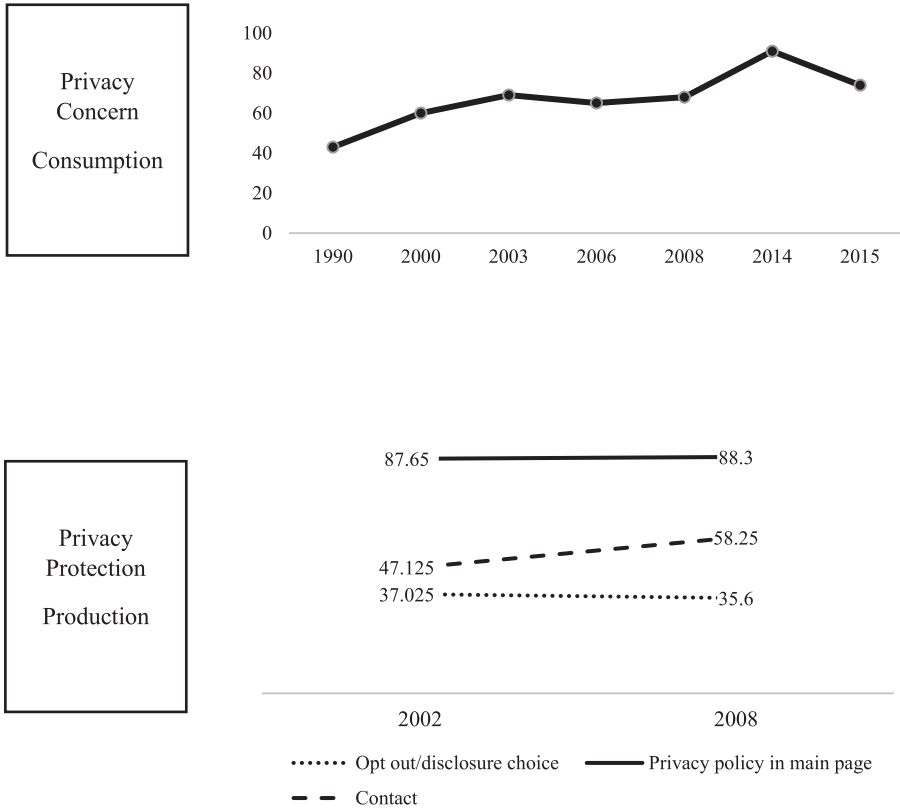

Figure 2.6. Marketplace Contrast between Privacy Consumption and Production. *Note*: For privacy protection, percentage of the sampled websites (Park 2008 and LaRose 2002) shown. See Appendix A for various sources for privacy concern.

We can gain further insight from analyses by the New America Foundation (2017), which offer the most comprehensive evidence available regarding privacy in digital marketplaces. The foundation's two-year wave survey rated privacy protection at leading companies (Google, Facebook, Yahoo, Apple, Samsung, AT&T, etc.) by conducting an 18-item assessment covering privacy disclosure, security, policy transparency, and other areas.

Table 2.2 shows ratings for each company, the difference between the two waves, and the average scores. Two conclusions emerge from the data: first, the latest assessment shows that average ratings are barely 50 out of 100 points (the possible total score), with tiny gains in 2017. That is,

the inadequacy of privacy protection measures is evident, confirming an ongoing weakness. Second, these are among the largest private companies, whose digital services cover most digital activities: social networking, search engines, online directories, email services, and even Google Glass–like wearables. In other words, when we slice apart the vast digital universe so as to highlight only highly popular firms, the mismatch between public concern over privacy and its protection looks even more dire, given the vast array of digital products and services offered by these giant companies.

2.3. Conclusion: Privacy, Market Externalities, and Dual Responses to Privacy and Surveillance

If we look at one or two specific areas such as health or e-commerce, highly detailed pictures of different sectors may emerge (see Rains and Bosch 2009 for health-related websites in the United States in their provision of Fair Information Practice principles). Evidence presented in section 2.2, however, identifies widespread practices and points in a consistent direction. All available data indicate that the supply-demand mechanism of the marketplace does not benefit privacy protection. This seems paradoxical because it suggests that digital production responds only to the market demand for surveillance, not for privacy. It is the core of the thesis formulated in sections 2.1 and 2.2 that the difference between market responses to demands for privacy and surveillance is structural, precisely because advertising, data marketing, and related sectors are financed based on personal data—a mechanism intensified by economic pressures toward fragmentation, personalization, and holistic measurement of audiences.

TABLE 2.2 Status of Privacy Protection: Major Technology Companies

	2017	2015	Change	Average score
Google	65	57	+8	61
MS	59	53	+6	56
Yahoo	56	52	+4	54
Twitter	53	51	+2	52
Facebook	49	36	+13	42.5
Apple	48	-	-	-
AT&T	47	52	–5	49.5
Samsung	30	-	-	-

Source: New America Foundation 2015, 2017.

Mine is not the first argument that identifies economics as the root cause of surveillance. The tension between market economics and other social and policy objectives has been debated among leading scholars in the following areas:

- Media diversity and concentration (Neuman 1991)
- Racially discriminatory marketing (Gandy 1998, 2012)
- Pollution (Hamilton 2001)
- Violence in television programming (Hamilton 2003)
- Gridlock over regulation of cable, broadcasting, and telephone services (Neuman et al. 1997)
- An FCC / public interest rationale that is at odds with economic efficiency (Bates 1993; Napoli 1999)

To this list we can add the friction between data surveillance and privacy protection, which is not readily aligned with the marketplace self-regulatory rationale. In this sense, one would characterize privacy as a market "externality" that can be represented as two contrasting trends—an oversupply of data surveillance, which may bring negative consequences, and an undersupply of privacy protection, which does not meet the market demand of those who are concerned about violation of their rights (see Bates 1993; Gandy 1998; Pickard 2013). Given the pointed conflict between surveillance and privacy, the market inclination toward privacy protection will be moderate at best. Digital media producers driven by AI-based market competition will find it difficult to fit privacy into their algorithm-based products and services. As in the cases of air pollution, concentrated media ownership, and violence in media, market profitability wins out against other social objectives that don't immediately impact corporate bottom lines (Hamilton 2001; Pickard 2013; cf. Stiglitz 2003).

Whatever one's perspective on privacy or surveillance, one obvious fact is beyond contention: digital platforms exist to be profitable. In fact, from the perspective of digital media producers and marketers, it is not only rational but also moral to maximize their efficiency and profitability. The business model that is optimized to maximize efficiency will appreciate audience-use valuations derived from institutional surveillance. The unintended consequences of these structural arrangements, however, can be troubling in the absence of regulatory intervention. Granted that there is no practical way for a person to avoid commercial websites and digi-

tal devices produced by private firms, George Orwell's description of Big Brother is not far our current realities. The matter is made more complicated because private companies, like Amazon, provide contractual data services for local and national governments, blurring private and public sectors where citizens resist the power of institutions.

In chapter 3, we will look at privacy from the perspective of ordinary people, to gain insights into how a person can resist or succumb to data surveillance. As it turns out, human agency is not so straightforward.

CHAPTER 3

A Perspective on Individuals

> People know what they do; frequently they know why they do what they do; but what they don't know is what what they do does.
>
> —Michel Foucault, 2003

3.1. Social Psychology of Privacy

The social psychology of personal privacy provides us with a sense of the conceptual tools people need to help guard against unwarranted digital surveillance. This chapter asks: How do individuals cope with surveillance? Under what conditions can a person exercise privacy control? Do digital technologies pose such threats that individuals are left with no power of resistance?

From the perspective of George Orwell, the future of privacy appears ominous. Orwell did not foresee a personal power to hide, screen out, manipulate, or disguise one's identity. Rather, citizens in their private homes, watched by a telescreen connected to the control center, have no option to re-establish private boundaries. The lingering image from Orwell is the telescreen as a tool of surveillance and oppression, offering a totalitarian control over passive, inattentive, and powerless individuals. Edward Snowden's revelation of NSA surveillance revelation is the most telling incident in support of Orwell's pessimism. The US public had no clue they were under such vast surveillance. Nor did citizens have means to resist the power of the US government. It is a reality of our day-to-day communication that the gathering of digital data has vastly expanded the secret monitoring of the economic, political, and social behaviors of almost every citizen.

Nevertheless, one must understand the dynamics of people's agency—no matter how tenuous—if we intend to protect privacy (Acquisti 2004; Ac-

quisti et al. 2015; Tversky and Kahneman 1981). As much as new digital technologies threaten privacy, we can imagine individuals with the ability to protect themselves, find a wat to circumvent surveillance and draw a line between the public and the private. This would be a delicate task, the private and the public are on a continuum. The thesis of this chapter is that human beings may never again be as powerful as Goffman (1967) posited in his optimistic portrayal of dramaturgical dynamics. But they may not be powerless as Orwell imagined in response to Big Brother. This chapter constructs a complex picture of individual cognition and behavior and their limits, and attempts to diagnose the societal conditions conducive of active privacy control.[1]

3.1.1. Privacy of Helpless Users

Three fundamental arguments have been offered to characterize people's agency in the control of private-public boundaries. These arguments summarized in Table 3.1. Many of the studies in which they have been advanced were conducted to understand the attitudes and behavior of the public regarding privacy, and here we can develop a typology of ordinary responses to digital technologies, notably the internet. Taken together, these arguments from various branches of social psychology paint a comprehensive picture of new-media users and agencies in their readiness to deal with digital surveillance.

The concern argument

Public sentiment about technological invasions into personal privacy have had its ups and downs. A majority of the studies referred to in the preced-

TABLE 3.1 Three Arguments on Helpless Users

Argument	Thematic	Typical technology	Agency
Concern	Paradoxical	Internet, social media	Passive victim and object of surveillance
Willingness to trade	Gullible	Digital marketing, online shopping, social media	
Convenience	Lazy	Personalized Web 2.0 services, social media, mobile applications	

ing paragraph attempt to detect trends in public concern through national surveys.

These studies have painted a consistent picture of worried citizens; in the minds of most people, surveillance conjures up the fearsome totalitarian potential of government control. The internet, commercial service providers, online portal sites, e-commerce, and social media platforms now occupy the primary fears of the public (Anonymous 1998; boyd and Hargittai 2010; Elueze and Quan-Haase 2018 for a critical review; Madden 2014; Madden and Raine 2015). Survey findings identify stratification among users in terms of their levels of concern over surveillance. Some of the earlier privacy studies (Ackerman, Lorrie, and Reagle 1999; see also TRUSTe 2008) focused on public attitudes regarding surveillance of e-commerce shopping. Westin (1998, 2001) measured public concern at a more general level and identified a three-public typology:

1. Privacy absolutists and those with a high level of concern over surveillance
2. People with little or no concern over surveillance
3. Pragmatists, those in the middle between the two preceding stances.

More recently, Pew internet surveys (2010, 2014, 2015) have found rapidly growing levels of anxiety and awareness regarding cloud computing, social media advertising, and marketing surveillance, as more and more respondents have expressed concern about their digital interactions. There is also evidence (Madden 2014) that suggests more Americans have become concerned about government and business access to their personal data since Edward Snowden's revelation in 2014. This trend may be erasing Westin's tripartite model and heralding mounting concerns about big data.

It is important to note that these streams of privacy research often converge on puzzlement regarding individuals' failure to take protective action, which seems incongruent with their levels of concern (Brandimarte et al. 2012; Norberg et al. 2007; Park et al. 2012; Park and Chung 2017; Utz and Krämer 2009; Xu et al. 2011). The so-called privacy paradox, in which people do not behave according to their expressed level of concern, rests on a core of passive individual users. This is a pessimistic verdict on human agency, as the absence of a response to the perceived threat allows the unrestrained exertion of surveillance power.

The willingness for "trade-off" argument

People's willingness to permit their personal data to be collected is in direct opposition to the romanticized notion of the concerned, rational, and active individual. People constantly confront the decision to release personal data, and studies in this vein have explored how willing individual users are to trade their privacy for such enticements as free access to content, monetary rewards, discounts, or prizes (Awad and Krishnan 2006; Li et al. 2012; Olivero and Lunt 2004; Phelps et al. 2001). This line of research has painted an image of complacent users who are not only reckless but willful victims—not heeding the consequences of their actions, they care more about the immediate reward and are willing to take the loss of their privacy rights for other gains.

Several empirical studies have focused on online shopping (e-commerce). Phelps et al. (2001) classified the types of information that online users were willing to trade; they found that financial identifiers were at the low end of the spectrum, whereas people were more likely to provide demographic and lifestyle information. Olivero and Lunt (2004) explored social psychological issues related to the decision-making in this trade-off. They explored individuals' perceptions of personal data exchange, collection, and extraction, and discovered that awareness of risk was the key to triggering concern about privacy and protecting against thoughtless release of data. Studies by Acquisti (2004) and Acquisti and Grosslags (2005) argued that internet users, although genuinely anxious about surveillance (the cost), often knowingly succumbed to immediate gratifications (the benefit)—a contradiction between behavior and attitudes that demonstrates the difficulty of controlling privacy.

A recent focus on social media users turned our attention to the release of sensitive personal data through posting, sharing, and commenting (Poltash 2012). Critics are concerned that private boundaries have been erased by limitless and instant broadcasting of political leanings and other preferences, pictures, and the details of one's social circles. This line of argument deplores how gullible users are in recklessly revealing information that in the past was considered private. A study by Dwyer, Hiltz, and Passerini (2007) demonstrated that online relationships can develop even on websites where the levels of trust and privacy safeguards are weak, conjuring up an image of clueless, carefree individuals who are willing to forsake privacy to gain of other interpersonal goals. Of course sharing personal

stories is one way to establish intimate social relationships. However, the critics' concern is that social media users usually prioritize sharing, open boundaries, and social relationships over privacy.

The convenience argument

The convenience argument has been frequently neglected in the debate over privacy. Convenience is often not the product of a conscious choice. Like any other human bias, convenience can be an automatic criterion that people rely on to make a decision, often without conscious awareness (Kahneman and Egan 2011; Tversky and Kahneman 1981, 1986). Convenience can come into play in such behaviors as heuristic inaction, succumbing to comfort, or being a couch potato. We can see this as a new-media version of the "lazy audience" thesis, which stipulates that instances of active cognitive involvement are relatively rate, for example, in choosing to watch a particular television program; instead most viewers follow a habitual, inattentive routine of mass-media reception (cf. Dutton and Blank 2015; Neuman 1991). Many pundits and policymakers warn us against such inaction, a concern encouraged by our commonsense observation of social media behaviors on Facebook, Twitter, Instagram, and the like. Some studies have documented the psychology of convenience and the potential risk, lamenting the lack of public attention to the effect of convenience on the important matter of privacy rights.

A body of empirical evidence shows that online users rarely bother to read privacy policy statements in their web browsing (Cairncross 1997; Culnan and Armstrong 1999). A policy report by Turow et al. (Turow et al. 2012; Turow, Hennessey, and Draper 2015) also showed that US consumers often succumb to the daily habitual routine of data release instead of making conscious choices because most of them feel it is futile to reject surveillance. This research is a refined explanation of the willingness to trade away privacy; in this case, the fundamental psychological underpinning is taken into account—the desire to be free from the cognitive burden of active engagement in favor of a habitual routine of preferred convenience (Acquisti et al. 2015; Neuman 1991; Neuman et al. 2011; Park, under review).

In a similar vein, the latest studies based on the social psychology literature show the role of optimistic bias, by which people underestimate the risk of privacy violation for themselves while they overestimate the risk for others (see Cho et al. 2010). Similarly, the "third person" effect (Chen

2018; Dienlin and Metzger 2016 for the privacy calculus model) rings true for decision-making related to privacy protection. Dienlin and Metzger (2016) found that people tended to evaluate themselves as having a lower risk of surveillance than others and more capable of dealing with privacy threats. These rationales are psychological justification for individuals to choose the convenience of inaction when it comes to protecting their own privacy—a lazy but convenient habit of allowing the release of their personal data.

3.1.2. A More Balanced Argument

We have so far created an image of passive, gullible, and lazy individuals who routinely resign themselves to the convenient habit of inaction. Empirical evidence documenting this behavior has accumulated over the years, and it would be naive to dismiss arguments about laziness that are based on this evidence. On the flip side, the image of powerful surveillance persists, against which individual agency remains critically weak. Nevertheless, there are two conclusions that one should avoid. The first is that people have no ability to resist digital surveillance and to retain control over their privacy. The second is that agency is purely a product of individual minds and rationality,[2] divorced societal context. Both of these conclusions would be misguided.

One of the central arguments in this chapter is that although people are demonstratively passive about privacy, the social psychology of their behavior is more complicated than a diagnosis of pure passivity suggests. After all, you get different results in flipping the same coin. People may not be either active or passive, either victim or villain, either gullible or clever. While people seek convenient inaction, trade personal data for rewards, and behave in ways that contradict their own concerns, they can also be attentive, careful, and systematic. A realistic assessment of individual agency related to privacy will recognize the copresence of passive and active responses to digital surveillance.

Cognitive ability and social stratification

Our thesis is that people, in spite of their cognitive limits, are not mindlessly active or passive, as their behaviors are grounded in preexisting societal contexts (Neuman 1986, 1991; Neuman et al. 2011). To put it dif-

ferently, people do not simply accept or reject surveillance; rather, their behavioral decisions are embedded in social influences and personal frames of reference and knowledge.

Pierre Bourdieu (1984, 1990) theorized the "habitus"—the resilient influence of social backgrounds in the development of cultural capital, such as knowledge of, tastes in, and appreciation of art, cuisine, fashion, and the like. His insights are applicable to the case of privacy when we understand individuals' agency in the mediation of digital surveillance technologies within societal constraints (Van Deursen, Helsper, and Van Dijk 2017). It is possible to identify a flaw in Bourdieu's pursuit of "objective"' social conditions (Bourdieu may well have neglected the subjective will of an individual psychology). However, the strength of his thesis lies in the acknowledgment that individual behavior and its determinants, such as knowledge and cognitive ability, cannot be isolated from the societal conditions that shape behavior, no matter how weak or robust a person's resistance may be.

That preexisting frames of knowledge play a critical role in one's privacy behavior should be noted (Neuman 1991; Pool 1983). People are not empty vessels without frames of reference. That is, they bring their own life experiences, understandings, and awarenesses, which are produced and reproduced in their engagement with the societal environment (Bourdieu 1984). In this way, Bourdieu (1984) recognized "objective" social conditions while highlighting the preexisting frames of knowledge that guide human behavior.

Note the crucial function of knowledge in Goffman's (1967) account of presentational control as an inherent human activity. To perform effectively, an actor in a theater must understand implicit rules about how to present "selves." In the digital era, to perform effectively one must understand personal data flow—that is be aware of institutional surveillance systems and understand appropriate actions. Thus knowledge is a cognitive determinant of individual agency as it is socially constructed and conditioned. Put differently, knowledge is a critical enabler—if not a causal determinant—for informed citizen who are making choices about privacy and surveillance, with active control constrained by broad societal environments. Highlighting this critical determinant of privacy-related behaviors, we can reject a simplistic verdict on agency and see how responses to digital surveillance remain contingent on social determinants (Bourdieu 1984, 1990).

The critical point is that one's cognitive power is not the inherent feature of an individual mind, but the systematic outcome of socialization—which can be measured through social demographic backgrounds. Table 3.2 illustrates the socialization model that conceptualizes how societal-level difference trickles down to individual-level differences.

Thus, this chapter's central argument concerning individual agency is built on the following two propositions:

- People's behavior in protecting their privacy hinges on the extent of their knowledge, a cognitive principle that empowers them to take informed control of their digital identities (Hargittai and Hinnant 2008; Hargittai and Marwick 2016; Litt 2013; Litt and Hargittai 2014; Park and Chung 2018).
- The potential effect of knowledge remains derives from one's social demographic background; thus there may be disparities in the capacity to respond effectively to digital surveillance (Chen et al. 2018; DiMaggio et al. 2001; Hargittai 2002, 2006; Park and Yang 2017; Pearce and Rice 2013; Neuman et al. 2011).

3.2. Empirical Evidence in Two Strands

Various data sets were used in this study to test these propositions and identify the patterns by which individuals resist digital surveillance. Nationally representative samples provide the best analytical tool because they detect variations in public knowledge and behavior at the aggregate-population level. To this end, we must synthesize findings from a data set based on probability sampling (Appendix B, for methodologies). Still,

TABLE 3.2 The Socialization Model of Cognitive Knowledge

Socialization	The process in which one learns, understands, and/or behaves in ways that are acceptable to the norms of specific communities and environments	**Societal level**
Sociodemographics	Variables of race, gender, education, age, and income	**Group level**
Cognitive agency	Knowledge, cognitive abilities, and/or frame of reference	**Individual level**

caution is warranted. First, not all of the samples analyzed in this section are strictly equivalent. In addition, measurements of privacy-related knowledge and behavior vary from one study to another. A longitudinal panel study best tests socialization effects over time because cross-sectional data cannot show that the findings are immune from random variations.

Granted these limitations, however, it is still possible to paint a holistic picture based on the underlying trends across different data sets. Despite constant changes in digital tools and the technological environment, the fundamental nature of people's behavioral and cognitive constraints does not vary. My goal is to capture how people's characteristics affect their ability to resist, appropriate, and respond to the power of surveillance in their use of the internet, social media, smartphones, and the like.

Proposition 1: The effect of knowledge

Table 3.3 shows the levels of surveillance awareness and understanding of privacy-related policies (combined scores) and the results from multivariate regression analyses concerning their effects on privacy behavior. The regression coefficients display the consistent and robust functions of knowledge in all dimensions of privacy behavior, which strongly supports our general expectations. Such tight covariation patterns also suggest an explicit behavioral principle—that is, one's cognitive frame of reference remains a critical determinant of privacy-related decision-making.

TABLE 3.3 The Effect of Knowledge on Privacy Behavior

	Knowledge			Behavior	
	M	*SD*	Reliability	Social beta coefficient	Technical beta coefficient
Surveillance awareness	4.73	2.40	.79 (8 items)	.32***	.27***
Policy understanding	1.96	1.86	.73 (7 items)	.29***	.19***
				M = 24.81 *SD* = 9.18	*M* = 13.12 *SD* = 5.18

Source: Knowledge Network data.

Note: Covariates not shown in regression coefficients. *** $p < .05$.

Here the knowledge items were coded as binary (correct = 1; incorrect = 0) and then combined as indexes. Behavior measured the extent to which individual internet users engaged in privacy control (see Appendix B for all the behavioral items in the two-dimensional indexes).

On the other hand, we also detected low levels of public knowledge about privacy-related matters (see Figure 3.1). People were not utterly naive, as they possessed a basic understanding of surveillance. More than 60 percent of the Knowledge Network (KN) participants correctly answered true-false questions about targeting of individuals and retention of their data. Still, some 40 percent of respondents could not correctly answer rudimentary questions about behavioral tracking, data collection, sharing, and tracking. The data show even lower levels of public awareness of privacy policies. Less than 23 percent of the participants gave the correct answer to five of the seven questions about privacy policies.

The results from analysis of bivariate relationships support the central proposition—the more knowledge a person has, the more actively the person is involved in control of privacy. Nevertheless, mass psychology may not demonstrate such bivariate simplicity. As discussed earlier in this chapter, key psychological factors such as a low level of concern over privacy and a willingness to trade it for compensation indicate a complicated scenario in which even the most concerned citizens may not make the effort to defend their privacy.

The effect of knowledge can be tested more closely. Notably, we can test whether the puzzling contradiction between a concern with privacy and careless behavior, for instance, may be explained by the lack of knowledge. We suspect that an individual's knowledge might prevent gullible behavior, such as mindlessly okaying the release of one's data release or trading privacy for small benefits. As we have seen, even those attentive to privacy sometimes compromise their privacy for reward or remain passive if they do not have enough knowledge to guide their actions (Baruh, Secinti, and Cemalcilar 2017; Neuman 1986, 1991; Pool 1983a).

Figure 3.2 shows the moderating effects of knowledge. Here a higher level of knowledge (both understanding of policy and awareness of surveillance) was positively associated with behavior to protect privacy. Put differently, control of privacy was greatest among those with a high level of knowledge and a high level of concern. Note the negative and significant main effect of privacy concern ($\beta = -.09$, $p < .05$), which shows a privacy paradox: that is, the concerned people were less likely to engage in privacy control. This finding implies that the knowledge is a catalyst for active control and resistance among ordinary people who otherwise may remain inactive despite their fear of surveillance.

Likewise, even when taking into account people's willingness to trade

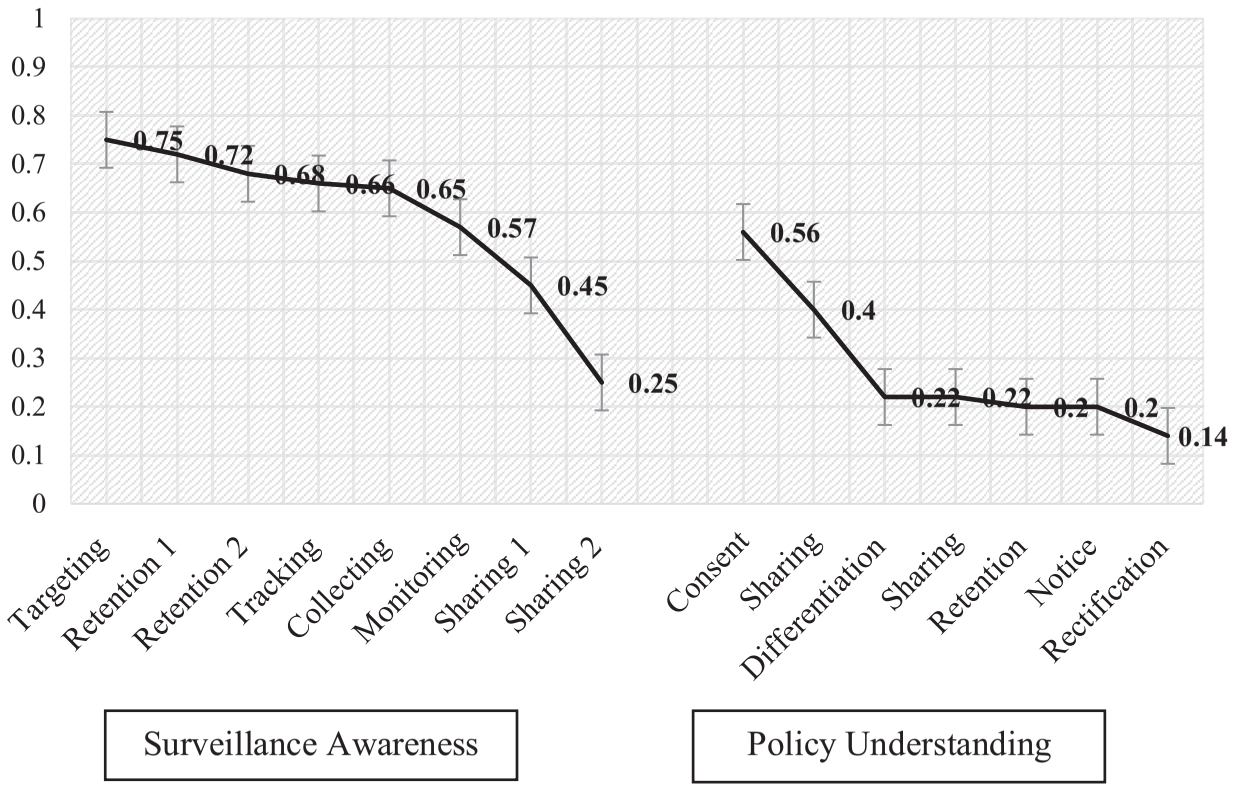

Figure 3.1. Distribution of Privacy Knowledge Items

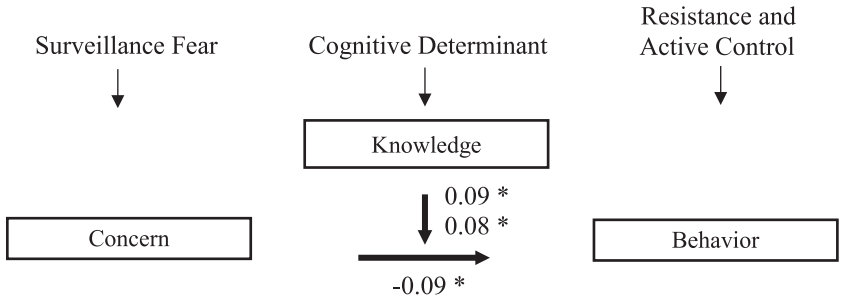

Figure 3.2. Moderating Effects of Privacy Knowledge. *Note*: Privacy concern is informational, with the behavior indicating the level of privacy control in technical dimensions (Appendix B). *$p < .05$. *Source*: Knowledge Network data.

off privacy, the behavioral effects of preexisting frames of knowledge was apparent. Among those who are concerned with privacy, people with high knowledge had increased levels of privacy control even when they had a high level of willingness to trade off privacy (policy understanding × concern × willingness: $\beta = .08$, $p < .05$; surveillance awareness × concern × willingness: $\beta = .15$, $p < .01$). In other words, the concerned people—however gullible they might be—were more likely to engage in active control when motivated by knowledge. This pattern, given that there was no significant relationship between the willingness to trade off privacy and privacy control (see Figure 3.3), suggests that most people rely on the convenient habit of inaction without the power of knowledge. Furthermore, knowledge widens the behavioral gap for those who are most concerned with privacy, and this gap appears to be increased by the willingness to trade off privacy because knowledge affects those with more willingness more than it affects those with less willingness to give up privacy for other gains.

Proposition 2: The incubating roles of social demographic backgrounds

A probability sample, which allows an analysis of social factors such as income, education, race, gender, and age (Van Deursen and Van Dijk 2011; Hargittai and Hinnant 2008; Neuman 1986; Rice and Katz 2003), is important in explaining the incubation of knowledge. If the links between the relevant sociodemographic variables and privacy knowledge prove significant, we will have evidence of the effects of socialization.

Figure 3.4 shows the effects of socialization, indicative of sociodemographics, on the development of privacy knowledge. Logic regression

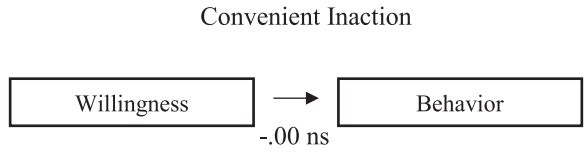

Figure 3.3. Direct Effect of Willingness for Privacy Trade-Off. *Note:* Willingness is the measure of the extent to which an individual is willing to trade personal data for rewards, with the behavior denoting privacy control in the technical dimension (Appendix B).

(odds ratio) and multiple OLS regression display the likelihood of having correct answers by respective variables, evidenced by the two populations: (1) adult US internet users in online use and (2) African American users of mobile devices (see Appendix B for all measures). First, adult internet users show consistent links between sociodemographics and knowledge. Gender is a consistent predictor, with males more likely to be able to answer questions on surveillance awareness and policy understanding. Education and race were also significant for surveillance awareness, as whites and people with higher education were more likely to give correct answers. Finally, younger people tended to score better than older people on questions about policy. Income was a significant predictor in both groups, suggesting the influence of one's financial status regardless of racial/ethnic background or the type of digital technology.

Interestingly, females were more likely to score better on mobile privacy knowledge, a difference from adult internet users. This may suggest the unique dynamics of digital communities, reflective of a social network structure in which women perform better or take leading roles (Feagin 2014). Importantly, social demographic differences appear to be replicated in a relatively young population that relies heavily on mobile devices (see qualitative observational data in Appendix B for cross-validation). In other words, when it comes to understanding of privacy-related issues, social influences persist even among a mobile-savvy population. Collectively, these findings suggest knowledge cannot be understood as purely individual ability.

With respect to the acquisition of knowledge, we can detect the influence of factors other than youth, such as the use of technologies, in the analytical model. We can make empirical observations about how one's appropriation of digital technology will interact with social demographic environments—that is, whether technological interventions such as the internet replicate the effects of sociodemographic factors or moderate their

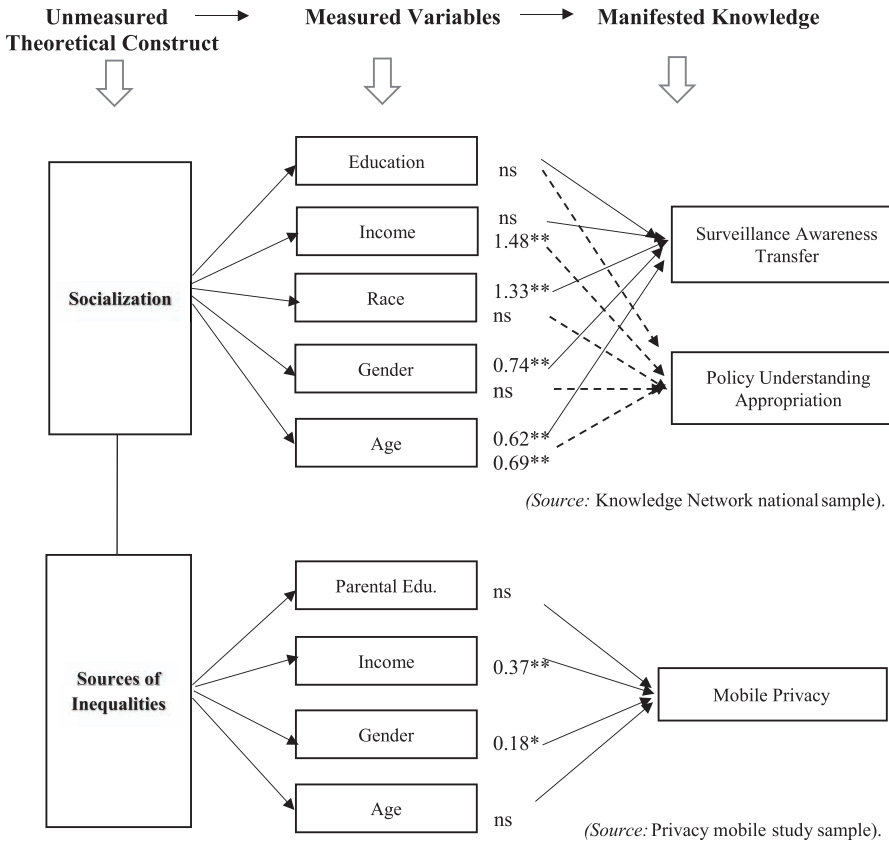

Figure 3.4. Acquisition of Privacy Knowledge according to Sociodemographic variables. *Note*: For KN sample, entries are odd ratios. The odds larger than 1 indicate the likelihood of the correct responses. Covariates (yearly experience; number of online accesses; daily use) are not shown in logistic regression. Policy understanding is an item asking about appropriation. Surveillance awareness is tone asking about transfer. Solid lines are for surveillance awareness; dotted ones are for policy understandings. For mobile sample, covariates (mobile familiarity; mobile access) are not shown in multivariate regression. Entries are standard coefficients. * $p < .05$; ** $p < .01$; ns = nonsignificant.

influences. If the intervention of digital technologies does not moderate the effects of socialization in the incubation of knowledge, variations in technological experience may simply serve as a means of reproducing, not eradicating, social demographic disparities in individual abilities to resist surveillance and control one's privacy (DiMaggio et al. 2001; Hargittai 2006; Sandvig et al. 2016).

Table 3.4 shows the results of the interaction between online experience and sociodemographics, using a sample of adult US internet users. Perhaps surprisingly, education and race, with their strong predictive powers, moderated the influence of internet experience. That is, the benefit of using the internet to gain privacy-related knowledge was most heavily concentrated among those with higher education and among whites. There was no significant interaction between online experience and policy understanding. This may suggest the extremely low variation in public understanding of policy, even when people have greater internet experience. Similarly, although the interaction with gender indicates that females gained more knowledge with increased internet experience, the main gender disparity related to knowledge acquisition remains significant in favor of males (for both surveillance awareness and policy understanding). Collectively, these results indicate the role of social demographic status in reproducing uneven cognitive power. In other words, the distribution of knowledge does not occur randomly; rather, social environments incubate one's cognitive capacity for resisting surveillance.

3.3. Conclusion: People's Cognitive Agency, Stratification, and Inequalities

The empirical evidence keeps returning us to a central theme in this chapter, which develops around two premises. The first is the function of an individual cognitive determinant that can translate passive, habitual inaction into active resistance and control of one's privacy. The second is the societal determinants that shape cognitive agency in digital environments. The collective thesis is that individual agency is a product of socialization,

TABLE 3.4 Interaction: Online Experience and Sociodemographics

Online yearly experience	Surveillance awareness transfer	Policy understanding appropriation
× Race	1.42**	1.03
× Education	1.30*	0.89
× Income	0.92	1.15
× Gender	1.26*	0.86
× Age	1.03	1.18

Source: Knowledge Network data.
* $p < .05$; ** $p < .01$

forged out of different social backgrounds (Bourdieu 1984; DiMaggio et al. 2001; Chen and Wellman 2005; Hargittai and Hinnant 2008). The potential effect of knowledge and cognitive ability is dependent on socialization, which exacerbates social disparities in people's capacities to respond effectively to digital surveillance.

Psychological factors present us with complex versions of human agency—people are concerned but passive, gullible, and willing to follow convenient routines of inaction. In other words, the model of psychological reasoning reviewed in this chapter, summarized Table 3 as three types of users of technology, helps us understand the processes of the human mind as it faces technological intrusion, making it possible to document inhibitory psychological mechanisms in individual resistance to surveillance. Individual cognitive ability, as indicated by the level of a person's knowledge, stands out as the fundamental variable determining privacy-surveillance behavior, and this reflects the influences of social environments that moderate the intervention of technology such as the internet.

By broadening this analysis into sociodemographic contextual variables, we can better understand how the distinctive characteristics of socialization shape one's cognitive power and become instilled in agency, which can be materialized in specific forms of knowledge. This chapter's findings are based on limited sets of cross-sectional data, yet the paradoxical nature of privacy behavior has been well documented in other studies (Hargittai and Marwick 2016). Moreover, the important role of knowledge, consistent with this chapter's thesis, has been identified and supported in recent work, including a meta-analysis by Baruh, Secinti, and Cemalcilar (2017), as well as earlier work on political knowledge (Neuman 1986). Fundamentally, cognitive stratification will mirror the resilience of sociodemographic influences, as it is expected to replicate the gap between the privileged and underprivileged (Buchi 2016; Chen et al. 2018; Hargittai 2006; Helsper 2017; Humphreys et al. 2012; Katz and Rice 2009; Marwick and boyd 2018; Neuman 1986, 1991; Park 2014; Pearce and Rice 2013).

If knowledge is the key determinant of behavioral differences, a systematic understanding of the root of cognitive power helps us reject a simplistic verdict on agency. That is, a realistic assessment of societal conditions is necessary to grasp why some individuals are more attentive to, knowledgeable about, and considerate of privacy issues than others, while most people are passive. After all, our view of the complex nature of people's agency must lie somewhere between pessimism and qualified optimism,

as we observe a general passive indecision and rare moments of active control, between justified skepticism regarding human rationality and faith in cognitive power.

Individuals who find themselves in supportive environments are likely to become aware of surveillance and know something about policies to protect privacy, enabling themselves to act. But those in socially sterile environments will less often develop knowledge and subject themselves to the tool of surveillance.

Making a case for such complexity remains a radical proposal.

CHAPTER 4

A Perspective on Policy Principles and Regulation of Data Flow

> What is responsible for the phenomenal development of the Internet? It certainly wasn't heavy-handed government regulation. Quite to the contrary: At the dawn of the commercial Internet, President Clinton and a Republican Congress agreed that it would be the policy of the United States "to preserve the vibrant and competitive free market that presently exists for the Internet . . . unfettered by Federal or State regulation."
>
> —Ajit Pai, FCC chairman, on the demise of the principle of net neutrality

4.1. The Logic of Minimal Policy Intervention in Digital Marketplaces

The 2001 Patriot Act that followed the September 11 attack, Edward Snowden's revelations about NSA surveillance in 2013, and the Facebook–Cambridge Analytica scandal in 2018 have several features in common:

- They precipitated a crisis of privacy and civil rights.
- They led to a public outcry against a growing level of digital surveillance in the United States.
- They blurred distinctions between the private and public sectors, as government-related surveillance programs run on popular consumer communication and digital platforms.
- They provoked calls for better measures to protect privacy and citizens' rights.

The Cambridge Analytica scandal has to do with election-campaign intelligences, the Snowden revelation is related to US government surveillance, while the Patriot Act is a legislative effort to allow the US government to intercept private communications.

What is often missing in public debates on these events is the fact that these were not technological crises, but political ones. In other words, they were not just problems of intensified surveillance, its "chilling" effects on public freedom, and manipulation of citizens' opinions and political decisions. Rather, they were the direct outcomes of conscious policy decisions made over time. Moreover, those "critical" events—immediately followed by intense public outrage, congressional hearings, and sensational media coverage—resulted in almost no changes of policy that protected privacy. This is a critical point made in this chapter, which argues that the future of privacy in an algorithmic and AI-entrenched world looks grim as long as political forces in the United States do not change an existing regulatory paradigm that does not address individual inaction (the subject of chapter 3) or the profit imperative (the subject of chapter 2).

This chapter critically examines the enduring legacy of US privacy policies so that I can recommend new policies in chapter 6. Suggested measures will be based on principles. What this chapter does in preparation for these recommendations is build a case for reform of personal data regulation (Neuman et al. 1997, 2016). Accordingly, it focuses on FTC policies and lays the foundation for suggested solutions, urging policymakers to revise, or even abandon, current self-regulatory policies. I will critically review privacy policies in the United States through the year 2017, when the FCC under the Trump administration began to relax its stance on net neutrality. This chapter will make a case that the lack of proactive policies is out of sync with the latest surveillance technologies based on algorithmic artificial intelligence, which are fundamental to such as Facebook AI, Amazon Alexa, and Google Glass–like wearables.

There is no inherent reason that the traditional privacy-related policies, which were developed under the auspice of misplaced marketplace ideals, cannot be mended. The lack of protective intervention in the United States reflects the political nature of policymaking within this marketplace legacy.

The marketplace metaphor and its legacy in US privacy policies

The marketplace of ideas is the most prominent metaphor in US communication policy (Napoli 1999, 2001; Neuman et al. 1997; Neuman

2016; Pool 1983a, 1983b). The notion of the marketplace of ideas is more than rhetoric; it serves as a fundamental basis for the shaping of US policy. The idea goes back to the seventeenth century, when John Locke pointed out that the attainment of truth is best achieved through free exchanges of ideas and information in the marketplace. In this view, government regulation is to be kept at a minimum to preserve the full functionality of a self-regulating marketplace (Dahlgren 2001; Horwitz 1991). Although one can interpret the metaphor in more than one way, a strictly economic interpretation has dominated in the United States, with a focus on the exchange of goods and services, often neglecting normative, nonmarket values and policy goals. Napoli (2001) elegantly noted,

> Within this interpretive approach, democratic theory does not factor into the definition and application of the marketplace of ideas concept. Instead, the marketplace is simply a place where goods are exchanged in accordance with the laws of supply and demand. According to this regulatory approach, the marketplace of ideas should be treated like any other product market and should function as such. (105)

In an affirmative sense, intervention via policies is a hindrance when a self-functioning marketplace best guarantees the sharing of diverse viewpoints and ultimately the truth. In most US communication policies, inaction is the direct consequence of this philosophical root. Policymakers have recognized the individual power with which rational citizens freely choose from among a wide range of options, fully informed in the modern marketplace as in the medieval town. Market institutions are assumed or theorized to encourage certain standards of action that fulfill democratic responsibilities in a self-governing society.[1]

One can easily understand why this principle of the marketplace of ideas is attractive to policymakers in the United States and is often translated into codified policies on protection of personal data and information. The individual-based solution of Notice and Consent/Choice that we looked at in chapter 2 is a prime example—once a person is notified of potential data risks and of rights to privacy, the burden of choice falls upon the user, who must decide what to do next. Instead of relying on a uniform solution that may soon be obsolete due to rapidly developing technologies, according to the marketplace rationale, the matter of privacy is best left to the discretion of individuals because they know what is the best for them, as they do in other markets.

Scholarly discussion has often neglected this neolibertarian philosophical root of privacy policy, which has exerted consistent influence on new-media policy in the United States (Neuman et al. 1997). Instead, transparency (Awad and Krishnan 2006), privacy by design (Langheinrich 2001), and accountability (Pearson and Charlesworth 2009) have occupied intellectual debates on data surveillance and on a better design for information privacy. There are viable alternatives to current US policy, for example, by prioritizing remedial measures and creating a legal responsibility for an organization that uses personal data. Legal scholars have analyzed legal precedents, for example in tort disputes in which a person files suit against a digital platform for breach of service terms and related violations. Yet these efforts, one would argue, treat symptoms, not causes. The fundamental cause of concern remains: adherence to the marketplace idea, by which supply and demand lead to institutional adjustment and the best available option eventually wins out.

Notice that from the perspective of the marketplace, disputes about privacy are reduced from a matter of citizens' rights to transactions between two rational actors in perfect power symmetry (the institution on the side of surveillance and the user on the side of privacy). But the rational actor theory loses its appeal given the feeble, gullible, and passive behavior individuals display, as reviewed in chapter 3, which opens them to bias and manipulation (Gandy and Nemorin 2018; Sapolsky 2004; Tversky and Kahneman 1981, 1984). Insights from game-theoretic studies (Guerra et al. 2003) also help us understand why it is futile to entrust privacy-related decisions entirely to individual discretion. Consider the following marketplace conditions:

1. Information is not equally shared between institutions and a person (who desires privacy).
2. A person's ability to cooperate, negotiate, or defect is hindered by the fact that the other actor (the institution) not only controls, but also monopolizes, the digital environment and its rules.
3. A third-party arbiter (judges, police, law enforcers, governments, etc.) that evaluates the violation of rules may also have an interest in gathering private data.

Julie Cohen correctly characterized faith in the marketplace as irrational (2012b). Cohen's thesis is that rational bargaining over privacy prefer-

ence is unlikely given that the marketplace has already tipped toward the disclosure of personal data as a condition of the consumer's entry into the market. This suggests a significant impediment to market equilibrium—the foundation of Notice and Consent/Choice as standards for protecting privacy. Even if we set aside the critical matter of trust that influences the disclosure of personal data (Dutton et al. 2005; Dutton and Shepherd 2006), too many variables and uncertainties distort digital markets for them to be able to establish parity and equilibrium between institutions and users in a rational game.

The regulation of personal data flow: FTC governance

The FTC does not have the explicit objective of maintaining a 'public interest standard.' In contrast, because broadcasting licensees, such as local television stations, rely on use of the public airwaves, the FCC regards them as subject to regulation in the public interest. Nevertheless, the FTC perceives privacy as a transaction to be decided by a consumer and a supplier in the marketplace. Among the illuminating FCC cases demonstrating its orientation are the following:

- Abolition of the fairness doctrine that promoted an informed public by mandating balanced viewpoints (Pickard 2018).
- Regulation of radio transmission for public safety (Hargittai 2001).
- The 2015 FCC net neutrality ruling, under which the Obama administration protected equal and fair access by the public (Lloyd 2010).

The FCC's role is to regulate communication and digital technologies in the public interest, whereas the FTC, as an enforcer of administrative rules, oversees privacy as a consumer issue.

The difference between the FCC and the FTC is best illuminated in terms of the regulation of political advertising on television. FCC restrictions imposed on political advertising, such as equal time access (i.e., competing candidates are guaranteed equal access to advertising spots, and broadcasters cannot deny their rights), are not applicable to digital platforms. As of 2018, social media microtargeting, illegal digital marketing and its regulation (e.g., by issuing fines for false claims), deceptive advertising, and data fraud and misuse fall under purview of the FTC; it investi-

gates them as business malpractice based on consumer complaints. This is an important point because the individual consumer protection—in lieu of an explicit public interest clause—take up significant space in the FTC's operationalization of the marketplace of ideas. The FTC, of course, is not the only government agency that exercises power over consumer privacy, but it enjoys the most influential administrative jurisdiction, as it oversees digital advertising—the bailiwick of Facebook, Google, Twitter, and Amazon.

The fundamental critique of these companies has been articulated by Napoli (2015) and Napoli and Caplan (2017). Social media organizations have argued that they are not news media companies; by doing so, commercial platforms circumvent public interest obligations and oversight by the FCC. Similarly, by reducing the issues of privacy to self-determination and rational decisions by an individual responding to institutional surveillance (Benette and Raab 2018), digital platforms can push aside their legal obligations and hide under the wing of the FTC, which does not enforce a public interest obligation as rigid as that of the FCC.

Figure 4.1 shows how faith in individual power, as opposed to defense of the public interest, is enshrined in hands-off policies in the United States. The x-axis in the figure indicates the degree to which a new-media policy favors ordinary citizens over business interests. The y-axis represents the extent to which a policy is oriented toward industry profit, as opposed to human rights like privacy. The resulting combination gives us a comparison between the United States and EU, showing contrasting policy visions that shape privacy protection (Dutton and Peltu 1996; Van Dijk 2012; Venturelli 2002). This comparison is not offered in order to compliment the EU's policies, which nurture social objectives under the public interest model (see Reidenberg 2001 for a comprehensive review), but to highlight the lingering legacies of market ideology, policy preferences, and collective beliefs about the size and the role of government (Neuman et al. 1997), which have an enduring impact on privacy policies in the United States.

In fact, a warning against optimism about the EU approach is warranted. Ambiguities about legal requirements and confusion about "the right to be forgotten" (implemented in 2012) and, more recently, General Data Protection Regulation (GDPA, introduced in 2016 and implemented in May 2018) show that the vision of a policy at its inception may not be realized in its consequences after implementation. A series of EU interventionist policies created anxieties among Silicon Valley

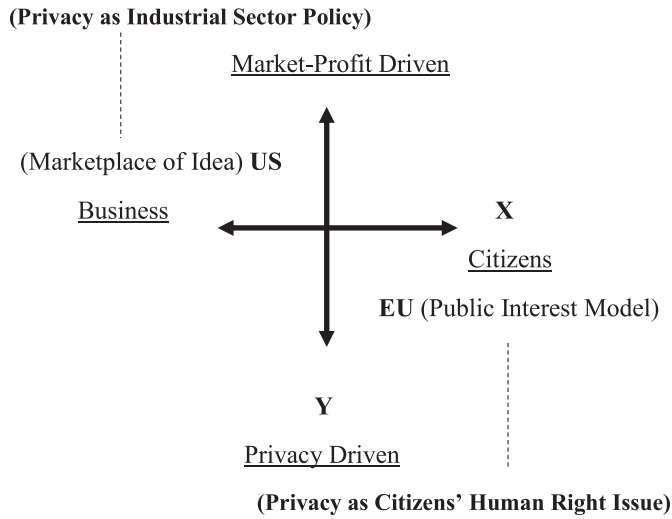

Figure 4.1. Privacy Policy Orientation. *Note*: X = Emphasis on well-being, as opposed to business interest.welling between business and citizens. Y = Policy preference between market-driven and privacy-driven. *Source*: Modified from Dutton and Peltu 1996.

companies—notably, Facebook and Google—which rushed to absorb the impact of rigid requirements from Brussels that, for instance, mandated the removal of particular personal histories from certain internet records so that they could not be found by a Google search. Even broader regulations were stipulated by the GDPA, such as the right to be informed about how personal data will be processed and the right to contest any automated decision-making by artificial intelligence. But the effects of these policies on the protection of personal information are yet to be established. One problem is that the top-down approach by the EU, which advocates a citizen's right to be ignored by a search engine, can hamper the public's right to be informed. In addition, the breadth of the GDPA's regulatory scope, some critics argue, makes it impossible for companies smaller than the giants to comply, while its burdensome requirements stifle start-ups. The lack of a clear, enforceable mechanism among EU member states, and a bureaucratic policymaking process that may not be nimble enough in its regulatory adjustment, cast doubt on the GDPA's success, even if we don't believe, as some observers do, that its true intent is to punish Silicon Valley.

My thesis is that whether it is the United States' self-regulation based on the marketplace of ideas or the EU's top-down commitment to the public interest model, viable policy options can be blindsided by their long-standing ideological commitments (Mueller 2009; Park 2009). US policymakers may have an unfounded faith in rational individual action that delegates hard decisions to the magic of the marketplace. This is problematic because the US policy process does not get support from an independent data commission, such as that in the EU. Nor has Congress empowered the FTC with regulatory authority beyond investigative power, which is in essence based on consumer complaints. The FTC or the Department of Commerce fills this vacuum of regulatory power, introducing comprehensive legislative standards and playing a minimal role as guardian of the market, thus fostering conditions in which the privacy exchange between institutions and individuals is left to the operations of the market.

Piecemeal regulatory evolution of privacy, surveillance, and personal data control over time

Table 4.1 delineates the evolution of self-regulatory policies that govern privacy, surveillance, and personal data flow in the United States. It is useful to think of the forming of privacy policies in two ways. First, it can be a response to a specific technological platform that entails a surveillance threat different from previous ones (first column in Table 4.1). Second, a new policy can be a political response to a national crisis related to surveillance, or to changes in the international policy regime (second and third columns).

A starting point is the tenuous basis of citizens' privacy rights—the Fourth Amendment, which prohibits unwarranted governmental intrusion upon a person. But its scope is limited to unlawful access and physical seizure by the state of personal possessions in face-to-face interactions. This tenuous constitutional ground leaves the commercial sector vulnerable, considering that the distinction between public and private sectors blurs with the rapid adoption of electronic devices through which governmental surveillance frequently occurs. In the 2013 NSA surveillance program, the fact that Google "pushed" data to the government, in addition to the government's "pulling" data from the Google database, indicates the government's heavy reliance on private databases. It is a flimsy basis for privacy and data protection tracing back to 1789, when the Fourth Amendment

was introduced as a legal guard against state surveillance. No matter how remote it may be, this is the best statement that the US Constitution offers related to surveillance—it still serves as the basis of protection against government-compelled intrusion, as well as warrantless, nonconsensual trespass and seizure of citizens' mobile devices and their data.

The absence of an explicit regulatory framework governing the protection of personal data is glaring in the period of mass media, from the start of the broadcasting industry in the 1930s to the advent of cable television in the early 1970s (Miller 1971). It took more than 180 years to introduce major US legislation that protects citizens against data surveillance. The 1974 Privacy Act—the most comprehensive protection against wiretapping—was enacted less as a response to the threat to privacy at the dawn of personal computer era than as a political response to the Watergate scandal (in which operatives connected to President Nixon's reelection campaign put political opponents under illegal surveillance). In 1970 the Fair Credit Reporting Act, which grants citizens a right to be informed of their credit record and to rectify errors, was an exceptional case in which the transition into the computer age triggered regulatory responses in data protection. Some touted the Fair Credit Reporting Act as a landmark law that devised concrete steps toward consumer protection at the federal level. With this act, they believed, the regulation of data collection, surveillance, and appropriation in the private sector would begin to take shape. Still, it is critical that the Fair Credit Reporting Act relied on Notice and Consent/Choice standards that have become principles of privacy protection, fundamentally based on rational individual action and complaints about violations of privacy.

This is not to say that there was a complete regulatory oblivion of privacy rights. Toward the end of the 20th century, a number of ad hoc regulatory solutions to the advancement of communication technologies were made. The 1988 Video Privacy Protection Act was enacted in response to the growth of home video rentals and was designed to prohibit the release of personal information regarding these rentals (Flaherty 1989). Similarly, Congress introduced the Electronic Communications Privacy Act in 1986 to restrict state access to transmissions of electronic data by computer, an expansion of the 1974 Privacy Act, which forbids wiretaps on telephone calls. Finally, the 1996 Health Insurance Portability and Accountability Act (HIPAA) added another layer of protection specific to the health industry by creating provisions limiting the use, transfer, and disclosure of

sensitive medical information. This was followed in 2010 by modification of the FTC's regulations to include electronic medical records (EMR). In 2015 President Obama directed the Department of Commerce to draft the Privacy Bill of Rights. This is a white paper that suggests that privacy should be understood as a human rights issue, recognizing privacy as such a right for the first time in US regulatory history.

TABLE 4.1 US Regulatory Responses to Privacy and Surveillance

Technological platform	Political crisis / notable event	International regime change	Policy response
Interpersonal			
Face-to-face interaction			1789: Fourth Amendment on search and seizure
Electronic and mass media			
Telephone	1934 Communications Act		1970 Fair Credit Reporting Act
Radio	1972 Watergate		1974 Privacy Act
Broadcast TV	HBO launched		
Cable TV	Sony Betamax sale		
Home video recorder			
PC and internet			
Personal computer	Apple Macintosh in 1984 Microsoft Windows in 1985	1980 OECD Fair Principles	1984 Cable Communications Policy Act 1986 Electronic Communications Privacy Act 1988 Video Privacy Protection Act
Internet	Netscape in 1994 Amazon in 1994 Yahoo in 1996 Google in 1998	EU 1996 Data Directive	1996 Health Insurance Portability and Accountability Act
Web 2.0			
Social media	FTC jurisdiction expanded into online 2001 9/11 Facebook in 2004 YouTube in 2005 Twitter in 2006	2004 Asia-Pacific Economic Cooperation Privacy Framework	2000 Safe harbor provision between EU and US 2000 Children Online Privacy Protection Act FTC FIPP[a] 2001 Patriot Act

TABLE 4.1—*Continued*

Technological platform	Political crisis / notable event	International regime change	Policy response
Mobile phone and smartphone	Apple iPhone launched in 2007		2010 FTC's Do Not Track List 2010 FTC inclusion of electronic health data in HIPAA
Web 3.0			
Internet of Things	2013 Google Glass 2013 Smart watches 2013 Edward Snowden NSA surveillance Amazon Alexa in 2014 Google Home in 2016	2013 EU Right to be Forgotten	2015 Obama framework: Privacy Bill of Rights, Department of Commerce
Algorithm-based microtargeting	Trump inaugurated in 2017 Facebook–Cambridge Analytica Scandal	2018 EU GDPA 2018 California Consumer Privacy Act	2017 FCC reversal of net neutrality ruling 2018 Safe harbor struck down by EU
Fourth Industrial Revolution[b]			
Artificial intelligence Robotics	-	-	-

Note: In this chronology, privacy policies/laws are at the federal level, not state. The California Consumer Privacy Act in 2018 is an exception with strong privacy protection.

[a] For an in-depth discussion and history of the FTC's Fair Information Practice principles, see Y. J. Park 2011.

[b] "Fourth Industrial Revolution" is an umbrella term that refers to the digital transition to AI, the Internet of Things, drones, etc. Klaus Schwab (2017) suggested the term. Without agreeing that these changes deserve to be classified as an industrial revolution, I adopt the term for convenience to demarcate how different time periods have dealt with regulatory challenges.

Nevertheless, we should considered these policy formulations carefully. First, the 1974 act and its update in the 1986 Electronic Communication Act concerned government data collection and access in the public sector. This creates a vast legal loophole in which the private sector is insulated from burdensome regulation in the use, collection, and retention of information. From a technological standpoint, this is an outmoded policy given that government surveillance occurs in such commercial digital platforms as Facebook, Yahoo, and Google. Many civil rights organizations, such as the Electronic Privacy Information Center, and civil rights activists, such as Hypponen (2013), argued that this unregulated environment made

possible NSA surveillance, with commercial platforms allowing backdoor access by the government to personal data stored in their databases. This massive surveillance was aided by the 2001 Patriot Act, which allowed sweeps of metadata and of the content of digital communications, including email, video and voice chat, videos, and photos, reversing the flimsy protections enacted in the previous era.

Second, by creating patchwork remedies that differ by sector, the government has allowed personal privacy to be "contextually neglected" (Nissenbaum 2001, 2004). That is, digitalized personal data tend to flow from one sector to another because data in a digital platform that is not specific to one industry sector can easily transfer to another sector. HIPAA, for instance, is effective only insofar as medical data originate from people's interaction with health professionals, but it does not govern health-related data derived from social media such as Facebook and Google, making microtargeting based on medical information and the transfer of related metadata to third parties perfectly legal. By the same token, the right to rectify one's personal credit history, narrowly defined under the Fair Credit Reporting Act, is not a right to reject or to be informed about discriminatory marketing and behavioral targeting based on digital traces related to a person's finances. Even the Children Online Privacy Protection Act, enacted in 2000 by Congress—the strictest provisions concerning personal data under FTC jurisdiction—applies only to websites or digital services that target children under 13, treating digital platforms that target teenagers or adult users as a domain outside of regulatory needs.

4.2. "Vast Wonderlands" of Privacy, Policies, and the Digital Ecosystem

Obama's Privacy Bill of Rights was drafted by the Department of Commerce, whose jurisdiction is the promotion of commerce and the free flow of information. This origin suggests that the US government's position on privacy will remain symbolic, observing it from the perspective of promoting US commerce. The Privacy Bill of Rights was only a policy white paper, far from a binding regulatory action, and it simply recognizes the need to update protections of privacy. It is noteworthy that the EU's "right to be forgotten" was implemented in 2013, and its higher standard of mandatory protections created a wave of regulatory responses around the globe. Certainly, the Privacy Bill of Rights was a symbolic step forward,

but with no real teeth it remains nothing more than a statement that the United States is aware of threats to privacy amid the advance of digital surveillance technologies.

As this chapter has argued, it is not difficult to identify a clear orientation in US privacy regulation up to the digital age, in which AI, algorithm-based devices, wearables, the Internet of Things, and Web 2.0 platforms such as Facebook, Google/YouTube, and Twitter gather and beam out personal data and make privacy more and more vulnerable to collection by various state and corporate entities. "A vast wasteland," Newton Minow, the FCC's chairman, lamented in 1961 about the programming on television. Though his remark was a normative statement about television, it reasserted governmental authority as an enforcer of the public interest in the broadcasting industry—the dominant information distributor of that day, just like today's digital technologies.

In 2018, forty-seven years later, the FTC commissioner Christine S. Wilson, in front of policymakers, industry leaders, lobbyists, and scholars at TPRC (Research Conference on Communications, Information and Internet Policy), struck a similar chord. While acknowledging a wave of digital transformation that is slowly pushing US policy away from a regime of Notice and Consent/Choice, she eloquently reminded the crowd that the FTC has two objectives to balance—consumer protection, on the one hand, and promotion of market competition, on the other.

Whether the two objectives must be separable as in a zero-sum game is debatable.[2] In any case, it is important to note the balanced approach described by the FTC commissioner:

- The promotion of market competition and the protection of privacy have the same regulative priority.
- Persistent faith is placed in the marketplace, the need for innovation, and business interests that thrive on personal data.

This is another example of US policymakers' understanding of consumer protection issues as deeply market based.

The FCC's reversal on net neutrality, with staunch support from the Trump administration, is not surprising in this context. It was consistent with the principle of an open internet and with individuals' negotiation of privacy within the market. The connection between net neutrality and privacy has received limited public discussion. By eliminating barriers for in-

ternet service providers (ISP) to discriminate against certain users and their access to content providers, the deregulatory move opened up uncharted "wonderlands" in which ISPs can collect and use personal data stored in their databases, and no one knows exactly how these personal data flow to third parties or are used—whether these parties are digital marketers, federal, state, and local governments, tax agencies, or the ISP's own subsidiaries. ISPs now can manipulate the content displayed and the speed at which users can access a site based on their personal data, and allowed to sell consumer information to advertisers and to create different prices for the goods offered. Whether and to what extent ISPs are actually practicing this option is a different question. What matters is the policy conditions under which the FCC, with its public interest values, is pushed aside and the FTC emerges as a de facto regulator of trafficking in personal data.

(Then) President Trump's personal animus toward Silicon Valley threw another odd twist into this dynamic. His support for the abolishment of the net neutrality principle, which hindered ISPs' behavioral targeting, effectively favors ISPs by placing communications companies such as Comcast and Verizon on par with Google, Facebook, and Amazon. This amounts to giving ISPs unfettered freedom to surveil and to collect and exploit personal data from home internet and cable subscribers. This policy is in line with feeble regulatory conditions that remain incapable of responding to data surveillance. The principle of minimal policy intervention still clings to the idea of the marketplace of ideas, despite its weak protection of civil rights.

PART III

Understanding the Future of AI and Its Challenge

CHAPTER 5

Ushering in the Era of Artificial Intelligence

> There will come a time when it isn't "They're spying on me through my phone" anymore. Eventually, it will be "My phone is spying on me."
>
> —Philip K. Dick, 1968

5.1. The Unregulated Industry of Algorithms and Artificial Intelligence (AI)

Decades ago, Ithiel de Sola Pool prophetically argued that the policymakers at the inception of new-media policy "cannot imagine the [technological] changes that lie ahead" (1983, 25; Napoli 1999, 2001; Neuman 1991; Neuman at al. 1997). Chapter 4 reviewed US privacy policies, which are examples supporting Pool's central thesis about the disjuncture between new technologies and policy. Pool contended that rapidly developing digital surveillance technologies would place unprecedented stress on longstanding beliefs about the self-regulating marketplace. As explained in chapter 4, the US self-regulatory regime is not simply a product of mindless policy inaction. Instead, US privacy policy (or the lack of it) is the outcome of a regulatory construct that did not evolve in a vacuum.

In this chapter, we translate these theses into more concrete insights by turning to the unregulated industry of personal data-based AI.[1] If most policy remedies designed to address the challenges posed by new technologies are based on past legacies, what might be the characteristics of surveillance threats that new technologies such as AI will introduce? Can we expect the emergence of a radically different paradigm based on AI in which a new system of protection addresses the challenges inherent in individual behavior patterns and institutional economic imperatives?

Sections 5.1 through 5.3 examine emerging algorithmic surveillance in the AI-based digital media industry. Section 5.1 reviews three dimensions of AI—its concentrated structure, the prisoner's dilemma logic of human interaction with AI, and the peculiarities of personal data as information—and investigates how they will contribute to increased surveillance. Section 5.2 illustrates new interactive dynamics between individuals and institutions through the cases of Amazon Alexa and Facebook / Cambridge Analytica. Section 5.3, about the normalization of AI, revisits the principles that shape digital surveillance technologies, which are fueled by institutional impulses that are currently under no regulatory oversight.

5.1.1. When AI and New Media Meet Old Regulation

New days of digital surveillance have arrived with AI.

Automated surveillance is everywhere. Producers of consumer goods and services, with intelligence about them and their purchasers having shifted from analog to AI-based algorithm processing, have an insatiable appetite for personal data, posing unprecedented risks to privacy. Amazon has already revolutionized shopping. Histories of individual purchases are permanent, and AI predicts what a shopper wants more accurately than the shopper does. Our social lives, once isolated geographically, are no longer secret, as people are constantly connected via social media platforms where their postings and comments, behavioral patterns, and personal profiles are collected, analyzed, and retained. In 2018, Google CEO Sundar Pichai reluctantly admitted that AI routinely scans people's Gmail messages and lets third-party advertisers microtarget individuals based on the content of these messages (McKinnon and MacMillan 2018). In this environment of ubiquitous surveillance, the government now has betters means by which to view citizens' lives, through obtaining access to a single digital platform. Facebook alone, for instance, can hold a vast array of traceable digital trails of personal records, effectively surpassing that of any national government database (Benette and Raab 2018; Solove 2001).

AI gobbles up personal data and surveils us automatically. Now that is a shift.

What hasn't shifted is old-fashioned regulation in the marketplace—or to be precise, the absence of regulation. Certainly, the existing regulatory regimes in the United States (and the rest of the world, for that matter)

have not anticipated these changes so as to effectively address their potential violation of citizens' rights, threats to privacy, and ubiquitous AI-based surveillance. Moreover, it is unknown how the patchwork system of data protection, developed for mass media and stand-alone computers through the early 1990s, will apply to networked AI and big-data algorithms. Was Ithiel de Sola Pool correct? How do the forces of production and consumption behind this digital transition sustain an environment that is conducive to surveillance and that supports marketplace itself over privacy? Will the old paradigm of nonregulation soon be challenged by AI-driven digital surveillance? If so, what policy tools do we need to build a bridge from the old to new paradigm?

The shaping of privacy and digital surveillance will never be separable from users (people) or system producers (institutions) (see Giddens 1983). Simply put, it takes two to tango. But we should dissect the nature of the interactive process in which individual people and institutions influence each other and produce an outcome from their interaction (Gandy and Nemorin 2018; Giddens 1983; Neuman 1991). Lawrence Lessig (2009) made an elegant point, that "[digital] code is a law"—how a technology is organized constrains people's behavior. Code in the context of surveillance is the programmed condition under which people and their behavioral data are monitored, collected, and appropriated under a set of algorithmic principles. The precise nature of these constraints varies from one software program or hardware item to another; however, Lessig's insight speaks to the fact that the industry of algorithmic surveillance, as in Facebook and Google, remains entirely free to set up, design, and modify the technological-architectural conditions that control people's behavior, their interaction, and the flow of their personal data.

Code for digital surveillance in personal data-based AI

The elements of code for digital surveillance, or the architectural condition, are portrayed in Figure 5.1. First, we see the vast flow of personal data from the top layer, where a person or a device user is located, to the bottom layer, where the physical backbone of digital infrastructure, namely the internet, is rooted. One might characterize this integrated flow of personal data across different levels of the ecosystem as the vertical integration of surveillance. Second, personal data are constantly pushed and pulled by a digital platform, whose databases can merge its own information about

users with information from other platforms or third parties. This can be understood as the horizontal integration of surveillance, in which personal data travel among different platforms and services in a business sector. If horizontal and vertical are combined, we have the unregulated data flow in our AI-based digital ecosystem.

It is not hard to see that this horizontal and vertical concentration of digital data ecosystems will deepen vulnerability to misuse of data, unwarranted access, security breaches, and exploitation or manipulation based on a person's detailed private record—to the extent that collection, appropriation, transfer, and algorithmic processing of personal data remain largely unregulated. For example, Netflix, the dominant subscription-based digital platform for viewing, can pass personal viewing data to in-house or third-party app developers who might better program automated video suggestions (vertical integration), while Amazon Prime, in a (hypothetical) strategic alliance, might want to share data with YouTube, such as demographic data and viewing patterns, in order to better customize Amazon Prime offerings and to increase Amazon's market share among digital video platforms (horizontal integration).

The structure of unregulated data industry indicates the following:

- The de facto status of the near monopoly in digital platforms whereby a single company controls most of the market shares (e.g., Facebook in social networks and Google in the search engine industry)
- The oligopoly of a very few manufacturers that dominate the Internet of Things (IoT), wearables and mobile and smartphone devices (e.g., Google Home, Alexa, Apple, and Samsung)
- The closely guarded proprietary standards of big-data cloud services (Amazon, Google, etc.)
- The virtual monopoly of ISPs (e.g., Verizon and Comcast) in most cities, where citizens have only limited choices over online access.

The net result, illustrated in Figure 5.1, is that the AI-based architectural code of digital surveillance, as structurally arranged by a few dominant institutional actors, Facebook, Amazon, Google, and so on, will gobble up a vast amount of personal data. Worse yet, an individual at the entry point of this tightly interconnected data surveillance ecosystem only has the option to "opt out"; that is, in the default setting data are automati-

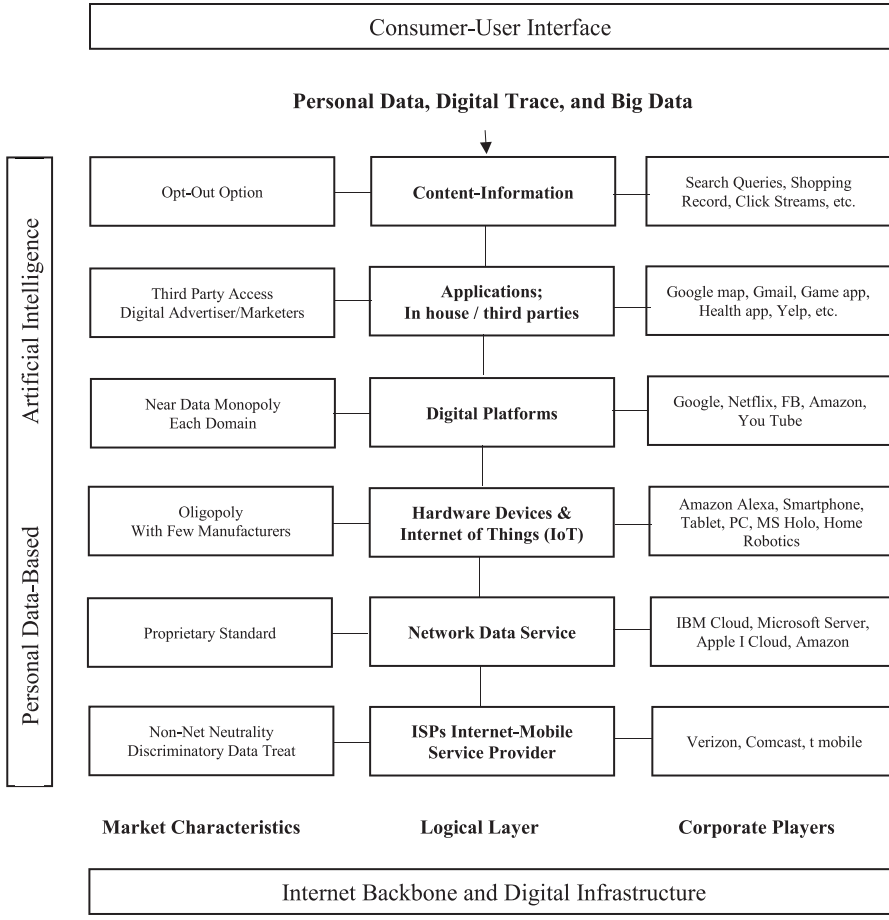

Figure 5.1. Personal Data Ecosystem in Vertical and Horizontal Concentration

cally collected by AI at the stage of consumption. Access by third parties, whether digital marketers or governments allowed to ride on a layer of digital networks, exacerbates this asymmetrical relationship in which the flow of personal data in a connected ecosystem in virtually unknown.

Here it is important to understand the concept of "quantified self/selves" (Cheney-Lippold 2017; Lupton 2013, 2016; see Van Dijck 2012 for "datafication"). This notion originally denoted self-tracked data or self-tracking behavior, especially in the context of medical diagnostics, but it also encompasses the idea that every aspect of a person's daily life can be

translated into a set of data points that are monitored and curated, often in real time. We can broaden this definition of "quantified self" to include the transition of surveillance into the AI-based big-data machine, which means the comprehensive AI-driven translation of people' lives and their social, economic, political, and cultural activities into quantified data, enabling real-time tracking, predictive analysis, and algorithmic reduction into certain values. Each layer of the industry sector adds its capacity to regulate and surveil individual behaviors as the entire ecosystem is optimized to collect data, reconstruct them, and create a 360-degree view of a person's identities (Danna and Gandy 2002):

- Financial worth (Turow 2003, 2005, 2017)
- Personal reputation (Solove 2001)
- Health status (Park and Chung 2017)
- Credit risk (Gandy 2012)
- Political standing (Park 2017)
- Housing qualification (Eubank 2016)

An important point about this code for digital surveillance is that AI is built to harvest as much personal data as possible so as to produce "intelligence" and to make decisions and recommendations—thus regulating behavior. Simply put, the principle is the more data, the better—the spurious logic of big data (boyd and Crawford 2012), but a principle adhered to in the AI and digital marketing industry. The unregulated AI industry by design is "data hungry," and its algorithms are ready to gobble up any personal data, even if it is of no use (Nissenbaum 2004, 2009; see Zimmer 2008). Pinpointing the actual harm or malicious intent of data surveillance is not essential. First, the violation of privacy has already occurred, as digital consumption presumes users' consent (whether explicit or implicit) to the status of being surveilled. Second, the uncertainties of how the collected data will be appropriated in a vast digital ecosystem warrant anxiety about the uses of the collected data.

This is Kafkaesque fear: AI, in every layer of surveillance, explicitly connected or not, is vacuuming all data, and the purpose of collection may be trivial at the time of collection but critical for later exploitation (Solove 2001). Even a single data point can make a person vulnerable to future manipulation. Rejection of digital services is impractical, for people doing so preclude their own participation in cultural, political, or economic ac-

tivities. The critical point is that in the layers of the vertically and horizontally concentrated data ecosystem, we are coded to lose a sense of privacy, data control, and, more fundamentally, the ability to determine how our identities are defined by algorithmic codes.

The burden of "nudge" and the appeal of "no nudge" against data surveillance

Oscar Gandy and Selena Nemorin's "Toward a Political Economy of Nudge" (2018) described a pattern in which omnipresent AI-based digitization, such as smart-city initiatives, can exacerbate the vulnerability of individuals, whose behavior is often hindered by cognitive constraints as well as structural conditions. Their argument is particularly helpful given their insight that privacy is a product of the mutual function of persons and institutions (Napoli 2015; Webster and Ksiazek 2012). They emphasize the asymmetrical power between individuals and institutions and the extent to which their interaction shapes the performance of unregulated markets, contributing to commercial exploitation of personal data in AI-entrenched smart cities. In this context, the "nudge," according to Gandy and Nemorin, is a necessary policy intervention without which individuals can be reduced to objects or "things" for the purpose of institutional algorithmic calculations (Sellar and Thompson 2016).

A reference to the prisoner's dilemma (Dutton et al. 2005; Neuman 1991; Hamilton 2000) can extend the argument by Gandy and Nemorin in an interesting way. Consider the diagram in Figure 5.2. A and B indicate respective values for an institutional actor (a digital AI platform) and an individual actor (a user or the user's data as an object of AI). When they decide to cooperate, both sides obtain certain values (A + B). In this case, cooperation means data submission or agreement that data will be used for AI. The AI platform obtains data, while users get (allegedly) better, customized, or personalized service, AI-automated recommendations, and access to the service. When they two actors do not trust each other and decide to defect, all lose values (–A–B). If one decides to cooperate but the other defects, we begin to have complications. To be more precise, if the defector is a user, we have A–B. The trouble is that when a user decides to defect, that often means de facto exclusion, not from a particular platform, but from that type of AI-based service, because a digital platform is often a monopoly.

This is different from a situation in which a digital AI platform decides to defect from data use (–A + B) or deviate from the way consent is explicitly or implicitly given. For instance, the AI might fail to predict accurately and recommend what a customer prefers. Or the AI does not deliver automatized service, or it may divulge personal data to third parties when exposed to an illegal data breach. The fact that the defection by a user will be forever known to the AI, whereas the defection by the AI can hardly be known to a user, attests to the asymmetrical nature of digital transaction (Masur 2018; Park et al. 2018; Sadowski and Pasquale 2015). In addition, the opportunity cost of defecting on the part of the AI is minimal, given that the user's data have been already obtained and that the monetary benefit of extracting the user's data might be realized not in the short term but in the long term, and will be based not on a single data point but on multiple data points.

The prisoner's dilemma assumes a purely homogenous and rational mindset with no psychological variation, and thus, the situation as presented above should be taken as a simplified version of the far more complex realities of human interaction with AI. That said, even in the most optimistic scenario, in which an individual user is fully aware of the risk of potential manipulation and surveillance in AI-based digital ecosystems, the cost and the benefit associated with the use of personal data will never be in equilibrium.

Notice that this is drastically different from a four-way traffic stop in which one person, by driving recklessly, can ruin a self-regulating traffic system, jeopardizing the safety of the other three drivers at the stop sign. In other words, the cost and the benefit of a possible defection will be clear to everyone involved, even without traffic police dictating drivers' behavior. Personal data market cannot enjoy the same dynamics. People not only have limited capacities (see Park 2013, 2018; Park and Jang 2014; Turow 2005), but also make decisions with incomplete information about the AI, while its algorithm has a disproportionately large amount of data about people. The four-way stop, where the expectations from the other drivers are rarely violated, is a rare human affair. In fact, it is unusual to see human rationality directing action even when we can see the cost of noncooperation, regardless of race, gender, income, or any other socioeconomic fact.

Notice in Figure 5.2 that even in the optimal cooperation of A + B, cost and benefit are not evenly distributed because the implicit cost for the

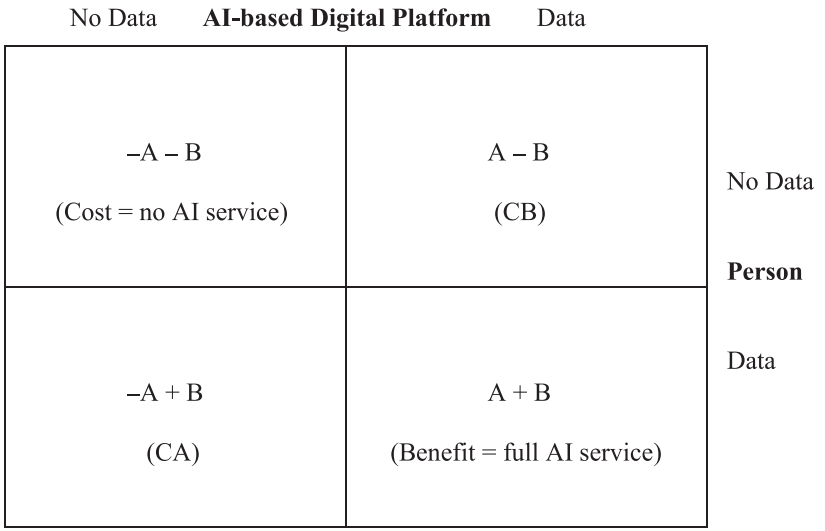

Figure 5.2. Data Submission (Person) and Use (AI) as a Prisoner's Dilemma. *Note*: A = value for digital platform, such as advertising, target ad, microcustomization, and marketing. B = value for a person, such as access and customized and automatic suggestions. CA = cost for AI, the opportunity to microtarget a person. CB = cost for person, AI-generated automation.

user participating in AI always remains—that is, loss of privacy, a status of surveillance forced on a user, or a sense of dignity that is reduced to sets of data points (Cheney-Lippold 2017; Lupton 2013, 2016; Van Dijck 2012). This is the dilemma. AI is fundamentally designed and coded for surveillance and this is never to disappear in its design AI. Thus, a nudge is warranted, a deliberate policy intervention designed to encourage a symmetrical interaction between AI and a user (Hamilton 2000; Gandy and Nemorin 2018; Neuman 1991; Neuman et al. 1997).

The peculiar characteristics of personal data as information

Another layer of complexity can be found in the unique characteristics of personal data as information. First, on top of the decreasing cost of data processing in general (Neuman 1991, 2016; Shapiro and Varian 1998), the collection, retention, transfer, and appropriation of personal data have become cheaper than ever. Moore's law in 1975 predicted that processing

capability would double roughly every 18 months, along with a decrease in cost per bit. As dramatic as this prediction was, it overestimated the current cost of data storage. The following conditions benefit the AI industry:

- The amount of storage capacity on a given chip doubles every 12 months.
- Optical transmission capacity doubles every nine months.
- The bandwidth of broadband and its infrastructure is expanding exponentially.

This trend of shrinking costs of personal data collection has critical consequences. It makes harvesting of data points far less costly than any human-to-human interaction-based information collection and thus makes digital data collection unthreatening to any business model based on an AI algorithm. Even when the massively surveilled data turn out to be erroneous or invalid, they cost almost nothing.

Second, the monetary value of processed data is far larger than that of raw personal data. In other words, the return on investment on personal data processing and appropriation is far greater than the sunk cost to obtain the data upfront. This is different from mass media, in which the first copy is enormously expensive, with high sunk fees, such as salaries and above-the-line production costs like scriptwriting, directing, and editing (Neuman 1991; Shapiro and Varian 1998; also see Couldry and Mejias 2018 for the abundance of personal data that are freely available for appropriation). With AI, the initial sunk cost, for example, in tracking a person's web browsing via bots, cookies, web beacons, and so on, is so low that any AI-generated value will likely be enough to recuperate the upfront investment.

Third, the increase in value will be exponential in both scale and scope once an initial value is created based on personal data. That is, the marginal cost of reproducing the value-added data profiles, by linking to and exploiting other databases (i.e., copying and transferring digital databases interconnected vertically and horizontally), will be almost zero. Data exploitation can continually deepen given a person's digital traces in an AI-driven ecosystem that is vastly interconnected and can capture virtually every facet of a person's life. The important lesson is that selling or transferring these comprehensive ranges of data to third parties like marketers, nonprofits, government agencies, and campaign fundraisers does not

diminish the value of personal data or preclude future uses by AI in its algorithmic processing of the already constructed databases.

Imagine a scenario in which person A and person B desire to eat the same apple. If person A eats the apple, person B cannot (exclusiveness) and the apple will be gone (rivalry). But personal data on that apple, on person A or person B, or on typical apple lovers will persist long after consumption of the apple as a physical good. Even one data point about where the apple was purchased may create tremendous value to the apple seller—as a third-party marketer can use it to create knowledge and databases—who may well benefit from AI-based targeting, whether it is addressed toward person A individually (microtargeting) or a segment of the population at the aggregate level.

The medical industry is the best example of this idea (Park and Shin 2020). Medical data collected from a patient can be used to develop a treatment or drugs after the care of a particular patient is over. A vast amount of knowledge constructed about the patient, the patient's diagnostic data, or genetic codes can be resold, refurnished, or aggregated into existing databases without losing any value created for the treatment of that patient. Furthermore, data points, such as behavioral traces suggesting sexual identity or political beliefs and affiliation, can be compiled to identify the type of medical conditions typical of, for instance, "affluent Democratic females" and may be used to target that group for similar medical treatment and drugs, as well as to create lifestyle-related advertising in mass-market offerings. The critical point is that the monetary payoff of the very first data gathered tends to be tremendous in the long term, because the personal data, as information, pull a disproportionate return from the initial moment of AI-based surveillance.

5.1.2. Individuals and Institutions in the Mutual Shaping of Algorithms, Big Data, and AI

Section 5.1 highlights three important points in understanding the rapid transition into personal data-based AI. First, the structural condition under which an individual person is situated by AI industries remains deeply concentrated and increasingly interconnected, as a few dominant institutional actors are in the best position to exploit the full dimensions of personal data. Second, in the performance of the unregulated AI industry,

the purported "rational" dynamics of the interaction between an individual and AI platforms, a prisoner's dilemma-like scenario, tend to situate a person as unable to exercise privacy and instead accept data surveillance as an entry condition of almost every digital participation (Lutz et al. 2018). Third, as people become subject to ubiquitous AI surveillance in every facet of their lives, personal data carry key characteristics (of nonrivalry and nonexclusivity), in that the potentially huge margin of return as a market product is propelling the intensification of massive personal data collection.

The fundamental point is that the transition to algorithmic AI will continue to reinforce the digital ecosystem that favors a surveillance-supporting marketplace interest over privacy protection. That is to say, of the AI-driven digital industry is unregulated, it is likely that the architectural code of AI-based surveillance will be further concentrated, designed to stifle a citizen's control of privacy, and optimized for data exploitation. This expectation suggests an urgent call for the nudge of regulatory intervention is required, beyond the familiar call for algorithmic transparency (Ananny and Crawford 2018; Pasquale 2015), by which we can open up a black box. It is important to note that even after a certain level of transparency is achieved, these emergent systems can use people's data in ways are disguised as consumer benefits or rewards. In this way, AI will function as a de facto gatekeeper that decides who participates in particular digital systems.

And this is the fundamental cause for alarm.

5.2. AI, Big Data, and Intelligent Machines

The following two cases illustrate the current context in which AI-based digital surveillance is shaped, illuminating the theses developed in section 5.1 about the rapid transition to an unregulated AI-based industry. This section takes a case study approach, based on data and insights generated from various scholarly, policy, and popular media accounts. The analysis draws on these data with the goal of resources to reconstruct existing AI-based surveillance conditions. Ultimately, this analysis exposes that the privacy protection in our AI-based digital environment has evolved in deficient ways.

This section argues that it is misleading to point a finger at a handful of companies—say, Facebook, Amazon, and Google—as evil-spirited

enterprises plotting invasions of privacy. The stances of Mark Zuckerberg and Sergey Brin on privacy may be a cause for public concern, but it is unproductive to single out their companies as the genesis of our disconcerting state of affairs, because such an episodic understanding suggests that reforming these companies or replacing their leaders would be remedies.

The AI system and its surveillance evolves not in a vacuum, but in the context of the limits of people's agency as they respond to the system of institutional practices and related interests (Barocas and Selbst 2016; Giddens 1983; Neuman 1991, 2016). These dynamics can be expressed as follows:

X1 = Individuals
X2 = AI-based digital media platform
X3 = External conditions, such as market competition, policy, and regulation, and so on
Y = Function of (X1, X2, X3, . . .) + X1 * X2, where Y = privacy-surveillance on continuum, with X1 * X2 as the interactive relationship between users and an AI-based platform.

As with the breakup of the telephone giant AT&T in 1982, the dismantling of Facebook by the US government would not eliminate the AI-based industry's hunger for personal data. Instead, the following cases, Amazon Alexa and Facebook–Cambridge Analytica, represent the emerging problems related to two business models of AI-based data collection and appropriation. These companies are the locus of long-term structural concerns, not a short-term meso-level organizational diagnosis, and thus they should be treated as symptoms of a bigger problem rather than its cause.

5.2.1. Emotional Recognition in Amazon Alexa, a Personal AI Device

We start with Amazon Alexa, a smart-home device equipped with AI, but it could be any personal robotic device such as Google Home or Google Glass–like wearables. The business model is a simple one: a consumer pays an upfront fee to purchase the device. But its operation is complex, entwined in vertically and horizontally concentrated data ecosystems, with the algorithmic capabilities to hyperpersonalize its services to the various needs of individuals (Danna and Gandy 2002; Gillespie 2014; see Rathi 2018).

Consider the scenario in which one's emotional states are monitored, surveilled, and captured as Alexa helps a person manage her daily chores. AI may be activated by unique voice or biometric data, such as fingerprints, facial or bodily posture, or odor. Alexa recognizes its owner's emotional swing toward anger, for instance, when her voice, whose sound waves are sampled and stored, shifts to a higher pitch. We can call this detection "emotional recognition," that is, a biometric AI-based application capable of discerning the pattern of a person's emotion by analyzing unique characteristics of the person's emotional fluctuations. These emotions can be a basis by which AI determines the level of casual interactions with the user and recommends where and when to shop, what to buy or do, and guides the user on daily chores. This is not a science fiction. By October 2018, Amazon had already filed the AI-based Alexa patent, in which the technical principle of emotional recognition is based on personal biometric data.

Figure 5.3 offers insights into how emotional recognition works in Alexa AI. First, AI built into Alexa can detect emotional traces, such as tone, pitch, and volume, as well as word choices, convert them into databases used to classify emerging patterns into negative or positive traits depending on the extent to which certain data deviate from the normality of a default baseline, and then translate the data to recognize a particular display of feelings—sadness or happiness, anger or enthusiasm, anxiety or excitement, and so on, whenever the biometric traces corresponding to each of the classified feelings emerge. Finally, these precise microscopic views of one's emotional state can be appropriated by its algorithm to match the product offering or related advertising with a precise data point that accommodates and targets a consumer at a given moment.

Numerous exploitations are possible. Yet one major concern is health-related personal data. AI-based personal devices like Alexa will be equipped with health-diagnostic applications, especially for the elderly, cancer patients, or people with terminal diseases. They will have an AI capacity to monitor, retain, and process health-related user data and emotional swings in real time while cross-indexing other medical data, such as information searches, drug consumption, or the frequency of visits to a hospital, to predict one's life expectancy, which might determine health insurance quotes. In another possibility, the location data collected in Alexa-like robotics can be fed by default into an algorithm that processes its AI sequence, used in conjunction with other biometric data, such as blood pressure or emotional patterns, in order to profile someone's depression or predict the

likelihood of pregnancy among certain types of users—or even to calculate the probability of criminality based on emotional traits. This AI technique amounts to a human doctor's forecast of mental illness by continuously tracking the signals of emotional upswings from wearable sensors and thus identifies signs of abnormality in a person's biometrics, even predicting when her emotions are about to dip (Abdullah and Choudhury 2018).

There is a reason for concern over these abilities. New stories report that Amazon is providing AI-based facial recognition services, known as Rekognition, to Chicago and other municipal police departments for the identification of potential criminals, telling us how deeply AI surveillance technologies are already engrained in people's daily lives (Dwoskin 2018), which not only encroach on their rights to privacy, but also use seemingly banal data to produce intelligence that classifies people into a "sort" or a data set organized according to corporate databases (Bowker and Star 1999; K. Crawford 2013).

Here our immediate concern is the "scalability" of personal data in a tightly integrated personal data-based AI system—that is, the ability of AI algorithmic power to classify a person's emotional patterns captured in a vast range of personal data. These data can be linked through multiple points of daily activities to constantly produce and reproduce intelligence about the person. The overlapping data points in different digital platforms can be connected via an Alexa-like personal device that monitors the user's home in real time. This will open up analytic machine-learning specifications for future behavioral patterns, which in turn can be used to "lock in" the user by keeping her in the same AI-based environment from which she starts her daily routines. And in this process, her data will be surveilled, collected, and appropriated within the propriety system tightly integrated by a Seattle-based private company.

Imagine an Amazon Alexa that can function as an anchoring point of one's digital activities, capable of linking emotional trails of personal data to the viewing records from Amazon Prime, subscription and traffic data from the Washington Post and its website (which Amazon acquired in 2014), and cultural tastes and brand preferences stored in the history of the Amazon shopping cart, while promoting consumption within Amazon's vastly interconnected platforms.

As a matter of a fact, Amazon is already doing this, as it opened the first ever AI-based physical grocery store, Amazon Go, in Seattle in 2018. In this store AI and facial recognition can identify the customer, know

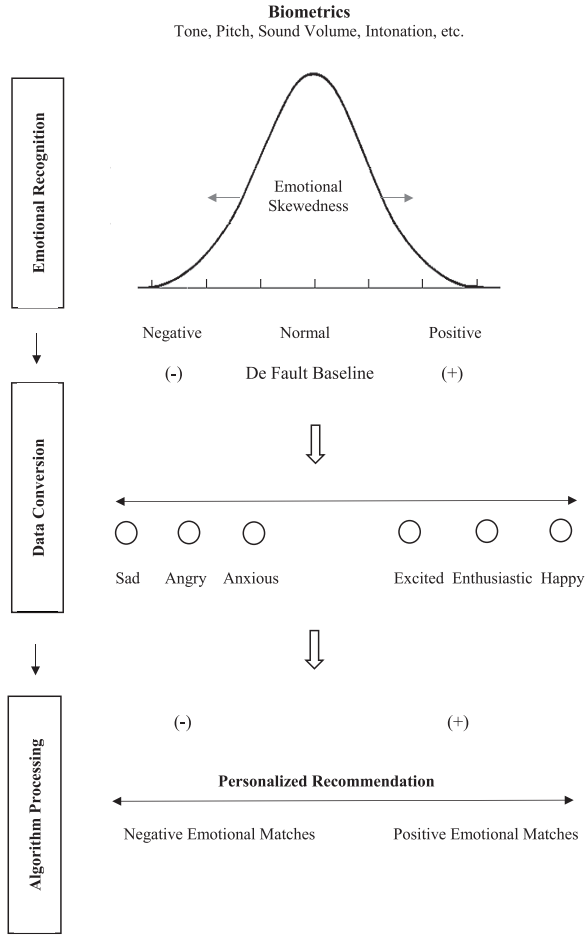

Figure 5.3. Emotional Microtargeting in Automated Shopping

whether she is a returning customer, what items she picks, what type of items she habitually buys, and so on. Then the customer can walk out of the store with the items she has selected because AI will charge the exact cost to her Amazon account (Gershgorn 2018). The implementation of AI-based facial recognition in major retail stores has been much anticipated. Yet the augmentation of emotional recognition is mindboggling: Alexa's emotional mapping of a person can be pulled from Amazon's deep databases at the right moment for "the best possible match" and be used to exploit someone's exuberant happiness to maximize the sale of particular products at Amazon Go, which will be able to capture billions of images, videos, and voices from a daily influx of unsuspecting shoppers.

In a more general sense, this AI-based digital mapping means that information about a citizen, associated data, and information of "actual or potential relevance to persons" (Couldry and Mejias 2018, 23)—through simple use of a personal device—can be captured within a closed proprietary algorithmic system (S. Crawford 2013). In this area Apple, despite CEO Tim Cook's plea for stronger privacy legislation in 2018, will be in a better position than Amazon to tap into personal data, given its vertical AI system, ranging from its data cloud, operating system, iPhone, smart watch, related wearables, and so on. Apple, just like Amazon, is in essence a personal data-processing company. No matter how much Apple claims to put its users in control of their own privacy, an undeniable fact is that massive data surveillance is fueling its innovation and thus its bottom line.

It is important to recall that the most primitive, but the most successful, business model that Amazon has built is its personalized list of recommended books, based on massive personal data sets collected out of shoppers' purchase histories. However, their rapid transition to AI poses a dilemma for companies like Amazon. Even when their business model is not solely built on personal data-based third-party advertising (through selling customers' "attention" to advertisers, so to speak) (Wu 2017), the Alexa and iPhone model will rely not only on selling an intelligent device or charging customers upfront service fees, but also on exploiting user data. This is a "walled garden" in which an exclusive set of personalized services are harnessed based on the privatization of personal data within an integrated AI system. The proprietary AI devices will tightly control a user's service entry and exit points, as their market success will hinge upon the digital surveillance of a user's private life for the purpose of personalization, matching, and eventually, business innovation.

5.2.2. Cambridge Analytica–Facebook and Microtargeting in the 2016 Presidential Campaign

On March 2018, a bombshell media report that the 2016 Trump campaign had access to millions of private users' data stored by Facebook, via the UK-based data consulting firm Cambridge Analytica, brought public outcry. This is a second case that epitomizes the transition to personal data-based AI. The outraged public and press reactions were well grounded. First, the amount of information held by Cambridge Analytica was massive, and given that Facebook has a de facto monopoly status, the

portion of the affected public was large. Second, many users realized how vulnerable they were to the algorithmic manipulation in the presidential election. Third, the widely publicized "fake news" of political misinformation in the "post-truth" era fueled public anger (Benkler et al. 2018), as did Mark Zuckerberg's lackluster response in US congressional hearings that worsened Facebook's image as a deceptive organization.

Despite heated criticism of Facebook, there is little understanding of the nature of the data surveillance that took place. Few legislative efforts or policy developments ensued, and there has been no apparent change in Facebook's approach to privacy since the scandal. This inaction suggests that the problem is structural rather than episodic. Most of Cambridge Analytica's use of personal data was legal. And because Facebook's open-platform model allows third-party access to users' data—unlike Amazon and Apple, whose AI data model is a closed proprietary system controlled by the parent company—we need to dissect Facebook's personal data ecosystem and how its AI surveils users, creates filter bubbles, and curtails or encourages exposure to selected political information (cf. Neuman 2016; Park, under review).

Figure 5.4 shows how Cambridge Analytica's data-based AI model may have implemented microtargeting to help Donald Trump's campaign. One of the first principles is classification into "targets" and "wastes" (i.e., the wrong population segment, not worth directing advertising to). This is a basic sorting according to the specific advertisers' needs—in this case, a political candidate. Imagine a composite score for an individual identified by AI as a target, a Facebook user who's Democrat-leaning in the 2016 election, but not quite sure of her support for Hilary Clinton.

The composite score of this individual can be calculated by pulling politically relevant demographic data points from massive Facebook databases: gender, age, rural or urban residence, ethnic background, income, consumption patterns, and so on. Analytica should grab these static data in addition to more direct political traces, such as comments on the election, postings, likes, and membership in Facebook groups, to build a user profile—a Democrat who may be swayed by damaging news regarding Hilary Clinton. Analytica could set its threshold at different levels; for example, only persons with data scores indicating a certain level of vulnerability would be classified as targets. It could revise its default algorithm as it collected more data to best estimate a target group. Interestingly, in 2016, Analytica conducted a Facebook user survey with questions about

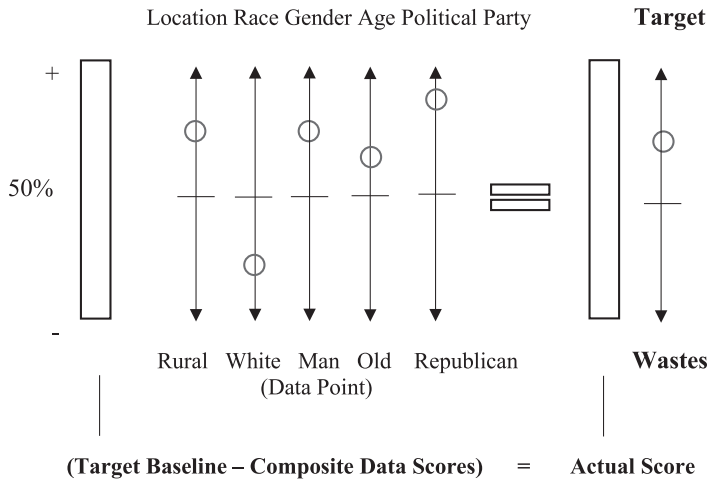

Figure 5.4. Political Microtargeting in Facebook–Cambridge Analytica

political beliefs, personality, and lifestyle, and used the results to create microtargeting strategies (Cadwalladr 2018). Think of adding a user's Facebook friend network, its demographic profiles and its political leanings, with information and newsfeeds to which the user has been frequently exposed, to these survey data, in order to identify a match or mismatch to the Trump campaign. The result is psychographic profiles of billions of Facebook users for whom target messages can be differentiated and sent out on a massive scale.

There are several important points to make about this strategy. First, the power of microtargeting is not the precision of the content of the message or its persuasive effect, but the ability of AI to classify users and to highlight or hide the information (or misinformation) in their digital environment (Neuman 2016; Hampton and Hargittai 2016). Second, participation in the survey by Analytica was taken as users' consent to have their data pulled from Facebook, since a Facebook account is de jure approval of data surveillance and appropriation by third parties. This is an important point because data collection has been justified in this way even when the personal data reaches third parties like Analytica. Third, Facebook's business model, by which revenues hinge upon third-party access and advertising, invites this type of microtargeting or data appropriation in the concentrated, loosely regulated data ecosystem surrounding Facebook. Fourth, the cost of pulling personal data is marginal, but the finan-

cial payoff to Facebook is huge, given that the same database can be sold to several third parties. By the same token, the Trump campaign's payments for Facebook's data are cost effective, considering that there exists no other single platform that can deliver more than two billion monthly users who are actively networking and willing to reveal personal details.

Figure 5.5 shows a macro view of AI-based microtargeting. The contrast between traditional mass media and the Facebook ecosystem is drastic. Let's start with media planning and buying in the context of broadcast television. The process is one-way and linear: advertisers buy commercial spots around a particular program—normally relying on advertising agencies that purchase and negotiate on behalf of their clients. The typical metric used in such a media campaign is audience rating; the larger audience, the more exposure for the commercial. In other words, a program's rating, no matter how imprecisely it portray audiences, is the best metric of the effectiveness of advertisements and is the currency that determines the prices of 30 or 60 seconds of advertising. (Here it is worth noting that the digital and mass-media ad markets are converging. It is increasingly rare to find an ad campaign solely based on traditional broadcasting. Even a local car dealership may buy spots on both local television and social media, which can help it tap into vastly connected personal networks.)

One clear difference in the Facebook model is that the "hit-and-miss" analytics in mass commercials are replaced by precise data-based AI metrics, behavioral traces of location, time, access, comments, likes, sharing, and so on. The effect of targeted ads or messages can be measured in real time. The Trump campaign had an opportunity to refine its targeted users based on the profiles they shared about themselves in Facebook, such as age, gender, relationship status, education, and type of work they do. What Cambridge Analytica did is overlay offline data from its user survey, measuring the effect of messages to develop more precisely targeted messages depending upon an individual's social networks—with all these data being fed back into the algorithm to better construct psychoprofiles and reduce any inefficiency in the microtargeting.

Markets in which personal data is exchanged between advertisers, political campaigners, and Facebook benefit the institutional entities involved. This is a structural issue, not an episodic one, because the data market fueled by AI works best at the cost of individual users' and citizens' right to privacy. Public anger toward Cambridge Analytica is understandable because it was the first publicly known case in which

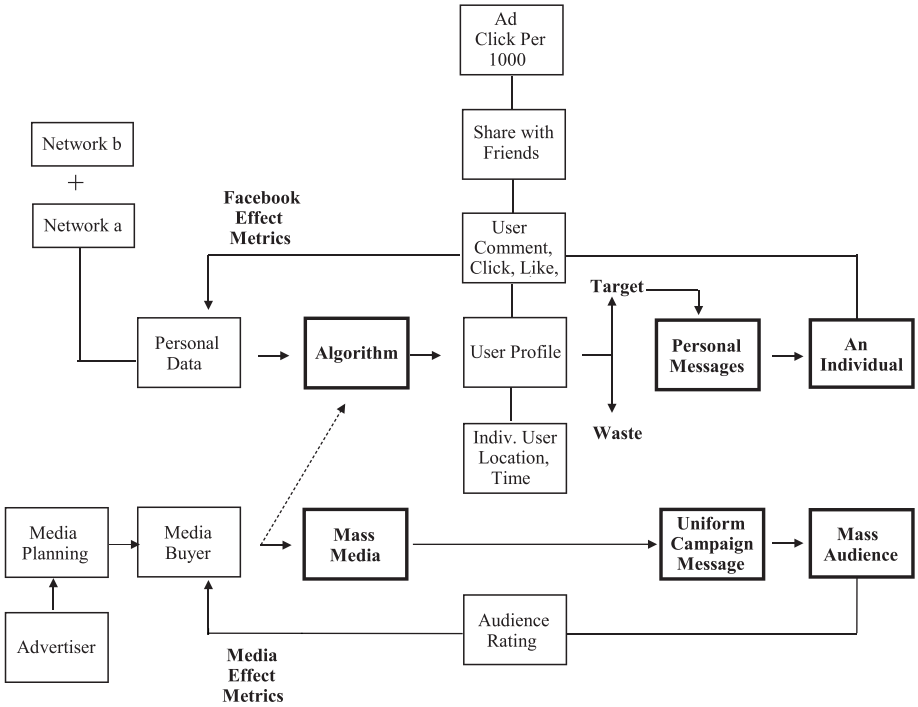

Figure 5.5. AI-Based Facebook Effect Metrics versus Mass-Media Effect Model.
Note: A dotted line indicates potential connection. For instance, media buyers often engage in microtargeting as well as a mass-media campaign.

AI-based manipulation happened in an election campaign. Yet such microtargeting should not be surprising given that this digital marketing technique has been implemented for decades with many commercial products, in the form of contextual ads (e.g., advertising of products related to one's search terms), display ads (e.g., a banner-type ad), and tailored ads (e.g., ads customized according to a user's preferences and sociodemographics), violating a user's privacy but often with that user's explicit or implicit consent (Turow 2005, 2017).

In other words, the trouble is that people did not knowingly permit their personal networks to be exploited by Facebook to organize information for the benefit of a third-party advertiser who happened to be a political candidate (Kramer et al. 2014; Meyer 2014). However, it is important to note that, under the First Amendment, political discourse enjoys

greater protection in the US legal system than commercial speech (Van Dijk 2012).

But AI blurs the distinction between political speech and commercial or marketing speech, and a self-regulatory policy does not see a hierarchical difference in the type of information that is manipulated, promoted, discouraged, or excluded by AI. This is remarkable because user choice is virtually nonexistent when it comes to Facebook, owing to its near-monopoly status in social media. A choice of channels or self-selection of certain information—let's say MSNBC over Fox—is not only unavailable to users of AI-based social media but predetermined by the person's private records and behavior related to her networks, friending, liking, commenting, clicking on a product ad, and so on.

In an AI-based digital environment, privacy and the extent of personal control over data surveillance determine one's exposure to diverse viewpoints—whether the microscopic AI-based information processing is political, cultural, medical, or merely commercial. The simple truth is that the refusal to let AI surveil us—that is, protecting our privacy—means exclusion from digital participation.

5.3. The Normalization of Data Surveillance in the AI-Based Ecosystem

The fundamental issue at stake is the "normalization" of digital surveillance in data-based AI systems. This does not simply mean the intensification of data collection since Web 1.0 era, nor does it refer to the transformation of interpersonal face-to-face surveillance, that has been prevalent in public interaction, into a digital form (Goffman 1967; Westin 1984, 2003).

Certainly, AI-based surveillance is different in type and intensity from earlier systems. But more fundamentally, the ubiquity of AI forces people to accept data surveillance as part of their normal daily routines—to the extent that citizens in all facets of life are surveilled and algorithmically defined (Cheney-Lippold 2017; Lupton 2013, 2016; Van Dijk 2012). The cases of Facebook–Cambridge Analytica and Amazon Alexa have demonstrated that digital surveillance—whether it is on the web, smartphones, or social media—is no longer out of the ordinary. Instead, AI-based societies, with an omnipresent, Big Brother–like AI, have reached a point where people see the lack of surveillance as "extraordinary" and non-AI environments as "broken" (Bowker and Starr 1999; see Jackson 2014; cf. Star 1999).

That is, the reality is already far closer to George Orwell's Big Brother than Erving Goffman's presentation of a self.

The continuous trend toward AI-based surveillance will lead to scholarly debates, policy discussion, and occasionally, public resentment—for instance, about yet another data breach of Facebook's billions of users in October 2018—less than six months after the Cambridge Analytica scandal broke out. However, as AI makes its way into the entirety of our digital experiences, the best prediction is that there will be less and less public shock and thus no change in the trend toward AI. Edward Snowden's revelation about the NSA PRISM program—a critical moment of government mass surveillance that involved every citizen's digital experience on the internet, mobile phone, emails, and so on—has largely vanished from public discourse (Lyon 2014; Park and Jang 2017 for discussion). In fact, we find little evidence to suggest that Snowden's revelation changed individual behavior to protect privacy. Nor did Facebook users migrate into other social network services after a series of Facebook data breaches.

This is not the same as claiming that people do not care about privacy (boyd et al. 2011; Hargittai and Marwick 2016; Park et al. 2012, 2018). Perhaps the opposite is true: people may be struggling with the complexity of the normalization of surveillance, in which the exposure of personal information is so deeply ingrained in mundane digital experience that one abandons hope of avoiding it.

Safia Noble (2018), in her poignant critique of the high-tech industry, chastised algorithms as a tool of oppression. Noble rejected a rosy optimism that search engines like Google could offer an equal playing field for different ideas, identities, and activities. Instead, algorithms often perpetuate bias against women of color and marginalized populations, as Google searches are designed based on people who are white and male. Similarly, Cheney-Lippold (2017; also Hintz, Dencik, and Wahl-Jorgensen 2017; Sandvig et al. 2016) offered a critical account of algorithms used for predictive policing, because they are often based on stereotypical labeling of people and other data like their web-surfing habits in order to infer their criminality. The thesis of this chapter supports these critics' assessment. Despites their importance in cases of political protest and expression like the Arab Spring, social media have exacerbated societal inequalities (see Park 2013, 2017; Park et al. 2018; Hargittai 2002, 2006). New forms of AI-based data collection that normalize surveillance can create further inequalities by using data to classify people as a type or a target under the

guise of bringing consumer benefits of personalization and customization (Bowker and Star 1999; K. Crawford 2013).

This book argues—fundamentally—that the focus should be on not just diagnosis and implications, but the structural conditions that induce this normalization of surveillance. Lessons from the transition from broadcasting to internet-based advertising are relevant. That is, it is naive to treat AI-based data collection simply as motivated by ill-conceived greed, an account in which the core problem lies in the impulsive pursuit of efficient microtargeting that new technologies enable.

The AI industry in its adaptation of a personal data-based model remains largely unregulated, which in turn provides structural reasons for AI digital platforms to be concentrated—a tendency fueled by the peculiar characteristics of personal data as information. Individual users, on the other hand, have the power to resist, appropriate, or even reinterpret data surveillance, but only within the constraints of a normalized institutional structure (Fischer 1994) that records, processes, and retains data about their digital activities. The fundamental dilemma, from the viewpoint of even the most able user, is that participation in an AI environment will always mean data surveillance.

We can summarize several lessons related to institutional and individual impulses with regards to AI-based surveillance. On the part of institutions, AI surveillance is a set of coded algorithmic calculations used to categorize collected data about a person's behavior for commercial, political, or cultural purposes (Cheney-Lippold 2017; K. Crawford 2012, 2013). That is, the augmentation of AI power is fueled by private records. Thus the surveilled data maximize the validity of automated decisions for better targeting, as in these cases:

- Facebook-Cambridge Analytica microtargeting millions of individuals for voter suppression across different platforms through digital ads, door-to-door interactions, phone calls, and mailers
- Amazon Alexa micromanaging a person's emotion to better construct precise knowledge about her.

5.3.1. Privacy Is Noise, according to Shannon and Weaver

Figure 5.6 maps out the sequence of AI decision-making, starting from (1) task identification, (2) input of personal data, (3) AI processing, (4)

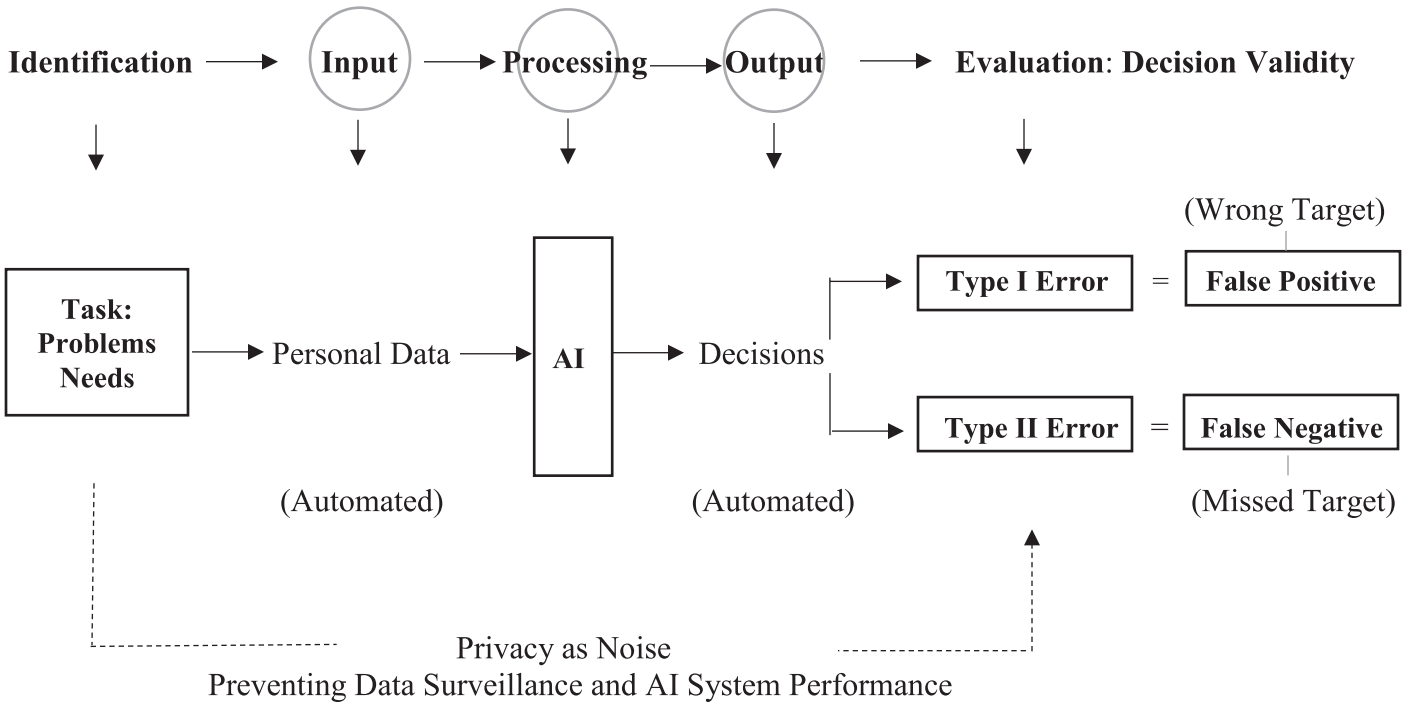

Figure 5.6. Sequence of AI Decision Error: Types I and II

its output, and (5) evaluation of decision validity. We can see how false negatives (type II errors), in which "wastes" are incorrectly identified (that is, wastes that should have been targeted), are serious errors that result in poor matchmaking (let's say, between sellers and buyers; advertisers and consumers; campaigners and voters) and thus have a negative effect on the bottom line. On the other hand, the cost of false positives (type I errors), in which the classification of a person as a target, or the targeted message, turns out to be wrong, is relatively low (Barton, Resnick, and Turner Lee 2018). From this imbalance, we can conclude that privacy or the lack of it is not the immediate concern of AI and its designers or owner. From the standpoint of economic efficiency, privacy is nothing but noise (Shannon and Weaver 1949) that hinders the maximum performance of AI and the fidelity of information it gathers because it prevents the collection of relevant data.

With less personal data input, Type II errors in AI are bound to increase. Again, the probabilistic assumption holds: the more data points, the more validity. Despite sampling bias (including a disproportionate number of white males in the baseline) (boyd and Crawford 2012; Gandy 2012; Hargittai 2018; Park et al. 2018), this makes sense economically. Take the example spam or bots. On the one hand, the indiscriminate barrage of spam offends privacy. On the other hand, for sellers this type of investment is cost-effective even when only one person clicks, given that the cost of sending duplicate spam messages or bots is almost zero (therefore, there less concern for type I errors).

Simply put, the respect for and protection of privacy (whether of an individual or an aggregate group) amount to noise from the standpoint of AI, whose primary goal is to reduce type II error. The fact that consumer nuisance and dissatisfaction, consequences of "not getting it right," rarely factor into the input (feedback) point in the design of AI says a lot about its priorities. Privacy concerns have never been part of the AI equation as an input data point. The central idea is that we must look at the economics behind the institutional inner workings of data surveillance.

The dilemma for Mark Zuckerberg, as well as for the AI industry in general, is that there seems to be no better business model than personal data-based AI, which is programmed to figure out the best selling point with automated instructions to read the private mind of a user. If AI is the best technology ever designed to augment human intelligence based on the collection of personal data, there will be no way to resist the force of

digital surveillance. This is the central reason to predict that institutional resistance to protecting privacy will be stronger than the forces in favor of protecting it, and as a result the conversion to better privacy will be tenuous (cf. Neuman 1991).

On the side of consumption, privacy protection against AI-based surveillance is also "noise" but for different social psychological reasons. Note the underpinning of individual decision-making in Figure 5.7, which shows how the level of reward (y), in the form of customization, access, recommendation, and so on, differs as a function of data submission (x) on the part of a user. That is, with a 0 (No) and 1 (Yes to data submission, surveillance, and use) binary decision, the steep pattern of reward is apparent to the user. This is to say there is the logistic function of being surveilled: let's say the submission of the first datum (1) is perfectly in line with one reward (1), for example, the initial access, with the submission of the second datum related to an automated book recommendation, and so on. It is just like a Google user allowing cookies, as the reward to the Yes decision will be greater than to No, with additional and continuous data tracking and surveillance promising another steep return of rewards, like automated search terms, search results customized according to the user location, and so on.

It is remarkable how the return of a reward serves as powerful psychological enticement to keep a user locked into the AI system, as constant rewarding of its automated decisions is lopsided (either yes or no) based on a single data submission, such as accepting cookies (Andrade et al. 2002; Jai and King 2016; Patil et al. 2012). Given the huge cognitive burden of taking additional measures associated with privacy, taking a reward in a split-second decision is an enormously attractive option (Acquisti et al. 2015). The AI-based precision of recommended Google search terms, relevant Amazon shopping lists, or favorite song suggestions in YouTube make the enticement of the reward even harder to resist.

This is particularly so because harms to privacy and the cost of data submission may not be immediately clear, but the reward is not only clear right away but also getting better in the short term, with even more use of personal details (see Kahneman and Egan 2011; Moon 2000; Park, under review; Tversky and Kahneman 1986). Starbucks's point reward system works exactly the same way: an individual is constantly invited to buy more coffee that is, more data points of when, where, how many, etc.), and finally, to reach a point of reward. AI can exaggerate this psychological

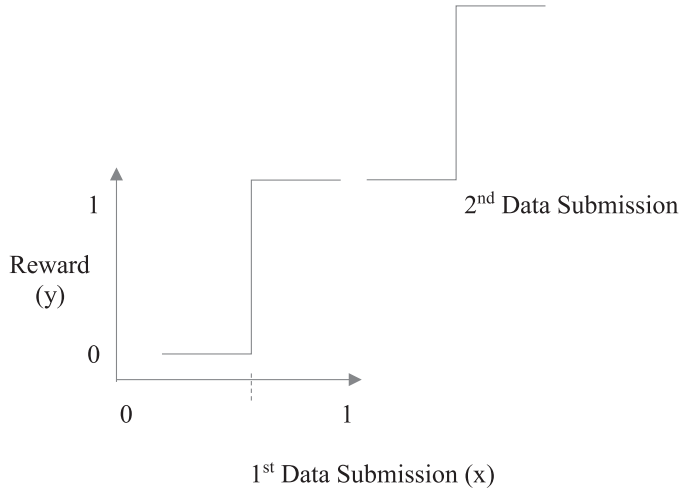

Figure 5.7. Logistic Growth Pattern in AI Return of Data

tendency, not just offering the same product (free coffee) but also more personalized offerings (free convenience), with seamlessly integrated digital services. In this context, an attempt to restore one's privacy hinders the fidelity of communication, in which the intended goal of digital consumption is not to protect privacy but to reap the best utility of a particular AI. In other words, in the AI-induced, highly automated, and risk-averse environment of digital routine and reward, privacy protection is noise even from the standpoint of an individual user.

Our convenience has come as a cost, then, the lost ability to define one's identity and selectively reveal it. That is, Erving Goffman's sense of privacy and active control in life's theater is now moot. This loss, engrained in a person's norma digital routine, is concerning. It highlights the very nature of digital consumption, in which the recursive digital system disallows the reshaping of identities, because users do not have the means to break away from surveillance or erase data (Giddens 1983; Napoli 2015). The fact is that there may be no reason to fight for privacy over rewards when someone's motivation is solely to protect privacy; restored privacy is often the privacy in only one platform when one's data has already flowed elsewhere.

"Privacy is noise in AI" is more than a metaphor, indeed.

Think of billions of Facebook users around the world, "voluntarily"

participating in its activities despite the fact that a majority of US consumers rated Facebook as the least trusted company when it comes to safeguarding personal data (K. Flaherty 2018). Even elite, knowledgeable users with a high level of concern for privacy succumb to the enticement of rewards, abandoning control over their privacy (Hargittai and Marwick 2016; Park et al. 2012; Park 2017, 2018). The ordinary consumer does not have a practical basis of knowledge on which to build a defense of privacy. Better privacy protection in the future will be even slower on the side of consumption than production.

5.3.2. Michel Foucault's Normalization

Michel Foucault (1975) in *Discipline and Punish* offered a critical insight that can be applied to the normalization of AI-based surveillance, in which new ways of collecting, retaining, and organizing personal data becomes routinely engrained in everyday life. To Foucault surveillance is a disciplinary device for exercising social control, as it rewards people conforming to its dominant rule, but punishes those who do not follow the emerging norm of proper conduct—thus reinforcing power.

The crucial power of AI lies in its organizing intelligence that can exclude those who do not conform to the routine norm of surveillance. What's critical to realize is that the normalization of digital surveillance is a product of the mutual contribution from institutions, motivated to design the optimal digital platform, and individuals, coerced to choose data submission "voluntarily," so as not to create "noise," and thereby avoid being excluded from digital participation. To put it bluntly, the differentials in the cost between data submission and nonsubmission and between being surveilled and refusal to be surveilled remain too large in the AI ecosystem to disrupt the normal digital experience to protect one's privacy—and create "noise."

Think of Google searches, a mundane activity. What Google's little white box invites us to do is insert honest words indicating what we need, think, want, what we desire to do, where we want to go, and, ultimately, who we are. We provide an enormous set of unstructured data about ourselves "voluntarily" and "honestly." In return, Google AI instantly rewards us with results that will in turn limit our choices and cognitive understandings, as our future behaviors will be effortlessly defined by Google AI. We

perceive data surveillance as a normal part of human-machine interaction when we consider an AI-driven digital platform like Facebook or Alexa just another social actor (Park, under review; Moon 2000), a place where we can develop intimate societal relationships through self-disclosure (Ho, Hancock, and Miner 2018; Jiang, Bazarova, and Hancock 2011). This interdependence, between otherwise distinct institutions and individuals, becomes normal, with the interplay of these two forces constituting the system of digital surveillance in the AI-based data ecosystem.

As illustrated in chapter 2, institutional motivation is banal, if not pure, in its pursuit of efficiency. Violation of privacy is not an intended harm but a by-product. Chapter 3 illustrates that people, in making decisions about privacy, are not simply passive or gullible in choosing inaction, but may be enticed by rewards in the form of free access, monetary compensation, customization, and so on. The matter is concerning given that individuals are enabled by knowledge but its acquisition is socially determined, as those with lower socioeconomic power may not have the knowledge resources sufficient to resist surveillance (Bourdieu 1984; DiMaggio et al. 2001).

The key point is that this mutual shaping by the two separate but interconnected forces of individuals and institutions jointly sustains the AI ecosystem and its surveillance practices (Giddens 1983; Neuman 1991). The AI-based ecosystem is structurally designed to encourage data surveillance—which in turn helps define and limit the scope and type of people's digital activities—and thus construct identities according to the maximum utility value determined by the system's needs (Gandy and Nemorin 2018; Park et al. 2018; Sandvig et al. 2016). As AI is set up to routinize the reward system to reduce any friction or noise, it is normal that institutions like Facebook, Google, and Amazon shape their respective digital environments in such a way that individuals' contribution to the maintenance of the existing AI system is solely based on surveillance and collection of their personal data.

In predicting the future of digital surveillance, we should note that this is the genesis of the problem. Since AI digital consumption as currently designed in most commercial platforms always correlates with surveillance, the intensified AI-driven world will rapidly transition away from the ideal of Erving Goffman—the control of oneself. The loss of privacy and of our ability to preserve it is not an episodic event staged by evil corporations. Instead, it reflects the personal data ecosystem's deep roots in economic imperative and mass psychology, which feed a self-serving system concerned

only for its own rights and survival (Michels 1915; Neuman 1991, 2016). Foucault's Panopticon, in which a guard in surveils every resident, is a prison. In our personal data-based AI ecosystem, where privacy is noise, the identity AI constructs for us is free.

Thus the evil is not Google, any other corporation, or the US government. Rather, it is systemic, lying in the bottom-line-driven profit imperative and the psychological nature of privacy behavior. These forces combine to reinforce the continued growth of AI-based surveillance. The threat of harmful results from surveillance is real because its output, as in the case of Google search results, might introduce, for example, misperceptions of women of color or discriminatory policing. Nevertheless, we must understand that the loss of privacy and other oppressive consequences of the application of algorithms are not intended effects.

Dannah boyd (2018) made this point in her sharp critique of how the algorithmic world perpetuates hate on the internet. Following boyd's line of reasoning, one can conclude that the banality of unintended evil is actually more problematic than intended evil. Structural banality makes surveillance a normalized routine of digital participation; any negative effects in defining who we are may be simply the result of differences calculated according to particular AIs, rather than any individual choice in how we present ourselves (Goffman 1967; May and Finch 2009). In other words, unintended harm may be more serious precisely because it makes more difficult any effort to differentiate the effects from the causes and the motivations of surveillance, and thus to repair the damage (see Gillespie 2014, 2018).

For this reason, in predicting the future of AI, I advance in this chapter the two-way model of surveillance between institutions and individual agencies. In the preceding chapters, I have argued that digital surveillance has emerged from a host of individual and institutional characteristics. Among them, a lack of morals has almost no place. Rather, corporate morals dominate US companies: profit maximization and return on investment. The immediate responsibility of Mark Zuckerberg, Jeff Bezos, Sergey Brin, Larry Page, and their peers is to their shareholders, their corporate board, and their employees. The general public pales in importance. This should not be surprising, given the lack of regulatory oversight to protect privacy and consumer well-being, and it will not change unless a public backlash threatens corporate images and bottom lines. We already know that such objectives as reducing pollution, encouraging diversity,

and combatting climate change have not been internalized in institutional behavior, despite loud calls for change. Moral arguments for the sake of normative merit alone do not offer sufficient rewards in the marketplace, no matter how important to humanity a particular objective—in the present case, privacy—may be.

This system invites regulatory intervention, which should be devised outside the two forces of individual agency and institutional imperative in the marketplace. As a matter of a fact, the central purpose of this book is to argue that the natural inclinations of people and institutions—either interactively or respectively—will *never* produce privacy rather than surveillance.

Jeff Bezos at his 2018 Amazon board meeting predicted the collapse of Amazon one day, saying that no company will survive a business cycle of 30-plus years (I. Hamilton 2018). His prediction provides a glimpse into how tech companies perceive the lifespan of their business models and try to prolong their own vitality.

AI-based digital platforms are offering a product like Coca-Cola, which will last quite some time, we can safely predict. We have witnessed the demise of AOL, Netscape, AltaVista, and soon Yahoo (as we used to know it) since the beginning of the commercial internet in the mid-1990s. So long as the rapid transformation of digital platforms into AI systems is based on "intelligence" deduced from algorithmic processing of personal data, AI-based companies like Amazon, Google, and Facebook cannot regulate themselves on behalf of privacy—not because they lack of morals, but because the institutional incentives that sustain them in the marketplace preclude allowing it to exist (Michels 1915; Neuman 1991, 2016).

PART IV

Conclusion

CHAPTER 6

Alternative Policy Principles, Options, and Recommendations

> The onus is on us to determine whether free societies in the twenty-first century will conduct electronic communication under the conditions of freedom established for the domain of print through centuries of struggle, or whether that great achievement will become lost in a confusion of new technologies.
>
> —Ithiel de Sola Pool, 1983

Numerous policy recommendations for better privacy protection have been devised over the last decades, yet most of the regulatory options can be classified into one of the four following categories:

- Privacy by design
- Privacy by law
- Privacy by market
- Privacy by norm

Countless legal experts, economists, information scientists, sociologists, and political scientists have presented somewhat disjointed but plausible policy propositions. Yet given the consistent market orientation of US privacy policies, as reviewed in chapter 4, skepticism about the proposed options is warranted.

This skepticism is not simply about aversion to the marketplace. Bureaucratic top-down approaches must also be avoided, because the staff in the EU Brussels office cannot execute all the technicalities of law all at once

and keep up with the advancement of AI algorithms, even though they have created a unified law. Even the 2018 EU GDPR, which purports to give citizens the power to access Facebook's databases and delete any transactional records it holds, cannot instruct citizens about the importance of privacy and how to exercise data control while participating in digital lives. The reality is that very few of us would want government to make privacy choices on our behalf—not to mention that governments themselves have an interest in surveilling citizens. We learned about these dangerous bureaucratic and totalitarian incentives from Nazi Germany (Budak and Rajh 2018; Flaherty 1989; Solove 2001) and more recently in the United States, from the NSA PRISM project, carried out under the guise of social order, security, and national interests (Park 2008; Park and Jang 2017).

The motto "Data are yours and therefore you have total control over them" may be an eye-catching aspect of Austrian lawyer Max Schrems's much-publicized crusade against Facebook and Google over alleged violation of privacy under the GDPR. This motto, however, only has the effect of making us feel good about privacy that we have already lost. Toppling one or two giant US companies is not a practical tactic for restoring privacy. No doubt action by regulators, which seems more likely in the United States in the aftermath of Cambridge Analytica, will jolt the unregulated marketplace into a new reality. Yet recall that Facebook, Amazon, and Google exercise the logic of the data marketplace; they are not the creators of the logic. Thus, my recommendation is not to think episodically, given the personal data ecosystem in which the flow of data is almost impossible to trace and the peculiar characteristics of personal data, which as information do not lose their value. Infinite future exploitations of data are possible with AI, as reviewed in chapter 5, section 5.1.

Simply translating the market-based solutions of the Web 1.0 era into the AI world would at best disguise industrial self-regulation. Privacy by design or engineering is intuitively attractive, but its focus on interface design or usability alone makes this option only a variant of the current industry standard of Notice and Consent/Choice. This is especially so because it will be up to individual industry sectors or a firm to adopt the design (Couldry and Turow 2014; Danna and Gandy 2002; Gandy 2012). Property-based solutions treating data as the person's private property, which have been proposed since the beginning of Web 1.0, remain rooted in an assumption that information is scarce and can be treated as a market product (thus, putting a price tag on data when a company buys

the private lives of citizens). But the abundance of personal data defies this logic, and the technical implausibility in both the private and public sectors makes it even more difficult to price the precise features of a person's data(for instance, where, when, how much, etc. in charging AI like Amazon's facial recognition technologies for their data uses). At best, privatizing personal data may correspond with an individual inclination toward the ownership of tangible market properties, in which the norm is fundamentally different from the norm of information churned out of infinite AI applications and processes.[1]

The point is not to repeat the cliché of a "balance" between market and government. We should be skeptical about an industry-driven hands-off approach, in which private firms are left to regulate themselves. In fact, the opposite is warranted: forceful intervention in the right market spaces should be made. To put it differently, the solution should not be variants of a market-based approach added to a governmental top-down policy. In principle, political forces and law, like market institutions, cannot keep up with the changing threats to privacy (Pool 1983a). What should be debated is the "quality" of intervention—namely, whether it is possible to create an intelligent combination of regulatory measures conditioning personal data-based AI environments that would not leave out privacy over the preference of data surveillance. This means "recoding" the digital data ecosystem so that (1) effective regulatory remedies can provide (2) mandatory design-engineering solutions (3) by which the marketplace can spur (4) norms for individual citizens to exercise privacy in their active management of public-private boundaries.

6.1. A New Paradigmatic Solution of "East Coast" and "West Coast" Codes

These principles are laid out in Figure 6.1. Given the skepticism about the marketplace that permeates this book, readers may be inclined to ask, "Why the market again?"

But that is precisely the point. Rather than an either-or debate—either market or government—over how to protect one's right to exercise control over personal data (see Neuman et al. 1997 for the falseness of the dichotomy of market and government), the discussion should be directed toward the type intervention that will be effective in the data marketplace newly governed by AI. In fact, even tort-based solutions, though problematic,

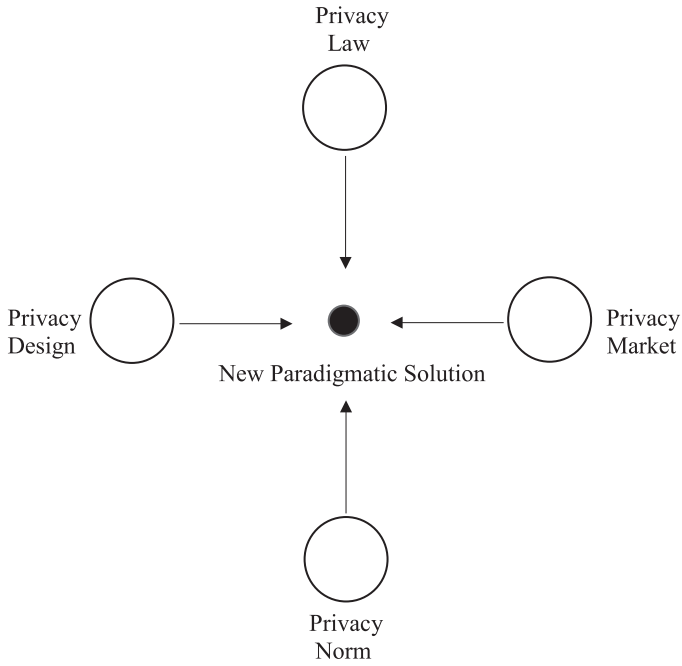

Figure 6.1. Regulatory Codes for Privacy over Surveillance. *Source*: modified from Lessig 2009.

can create enough pressure that a firm will have to worry about costly lawsuits (Posner 1981); this will be particularly true when the competition in marketplace is genuine, unlike the concentrated AI industry today. What's needed is regulatory oversight and broader policy principles that prevent the data marketplace from acting on its instinctive impulse of surveillance.

We can return to Lawrence Lessig's discussion of the clash between "West Coast" code, epitomized by technological innovation in Silicon Valley, and "East Coast" code, Washington, DC–created law and policy regulating individual options, freedom, and entrepreneurship. The binary distinction between East Coast and West Coast codes is an simplified one, but helpful to analytical and conceptual clarity. According to Lessig, East Coast code's public values of law and policy can result in better design of West Coast technical codes, as free-spirited West Coast is restrained for the benefit of public values and norms.

In this regard, Seattle-based Amazon's moving its HQ2 to Northern Virginia / Greater DC is intriguing. Positioning itself to curtail regulation,

Amazon provides a test of how a mix of the two codes will turn out. Lessig's original thesis in 1999 carried a tint of libertarian spirit, heralding the beginning of Web 1.0, in which an emphasis on balancing individual freedom with government supervision was common. For our purpose, Lessig's point can be reinterpreted to suggest that the marketplace does not need to be self-regulatory or unregulated to function at its best. In an AI world, the ideal digital marketplace may be the one that spurs the smart design of West Coast innovation tamed by the East Coast code—that is to say, an intelligent combination of regulations enabling normative options for individuals to control their privacy while still participating in digital activities.

In Erving Goffman's ideal of self-presentation (1967), privacy can exist even in public spheres when an individual is able, despite unwanted gazes of others, to control the presented self. Given how much of the stage of public presentation will be built by AI, Web 3.0, wearables, and other technologies in the future, the market's inclination toward surveillance will never cease. The task then is to recognize that policy's role will be to nudge private enterprises and institutions so that a person's digital participation is not subject to blanket surveillance. As Steve Vogel (1998) argued, the free market needs regulations—the effective functioning of the market depends on best managing its terms of freedom—that acknowledge there are needs the marketplace fulfills. Accordingly, this book insists on high-level principles that are flexible regardless of the specific technologies in operation, whose unique threats cannot be foreseen. More important, top-level regulations must not remain mired in ideological camps, privacy versus surveillance, or government versus market, in order to bring workable solutions to the table.

More specific policy proposals follow.

First, vertical integration[2] within and across new-media firms needs oversight by the FTC, which can be empowered by Congress to prevent the centralization of personal databases (i.e., keeping personal data in local platforms)—that is, a market solution. Second, the interface design of personal data-based AI applications, like social media platforms, smart-home devices, smartphones, and wearables, should be mandated to restrict third-party data access and the appropriation and retention of personal records—that is, a design solution.[3] Third, there should be federal, state, and local campaigns to promote public literacy on privacy and relevant digital skills—that is, a privacy norm solution. Finally, in the United States, Congress should empower the FTC to enact and enforce an updated opt-

in model regarding surveillance in personal data-based AI platforms. The last option should be stringently applied to sectors dealing with data concerning children, the elderly, and other vulnerable populations, as well as digital records related to personal finance, health, and other genome data—that is, a legal solution.

These combined measures are meant to suggest that there is no single remedy for algorithm-based digital surveillance. It is important to promote a set of foundational principles based on which policymakers can devise oversight and regulation to ease digital surveillance and the loss of our ability to control the flow of data and define our selves in AI platforms. The full scope of the Fair Information Practice principles (Notice, Choice, Access, Integrity, Enforcement) mentioned in chapter 2), though they are the most comprehensive working guidelines to date, remains difficult to incorporate in AI system design (Barocas and Selbst 2016). This concern about their practicality is in line with the skepticism of Ananny and Crawford (2018), who view these principles as a pragmatic limit, given users' bounded practices and knowledge in digital environments. "Pursuing views into a system's inner workings" (Ananny and Crawford 2018, 978; Gandy and Nemorin 2018) seems a remote solution, especially without a persistent effort to educate users about digital literacy (Park 2013) and awareness of how algorithms work (Sandvig et al. 2015, 2016).

Thus, when it comes to the specifics of AI-driven platforms, the macro-level solutions recommended above must be mixed with another layer of regulatory measures at a micro level. Particularly helpful is Lessig's premise that "code is law" (2009; Bowker and Star 1999; Neuman 2016), but here I mean the more literal sense of computer code in that algorithmic programming can be a powerful tool for regulating users' behavior as well as the operation of digital platforms. From this standpoint, it is telling that AI-based digital systems like Facebook, Amazon, or Google can be engineered or reverse-engineered for fulfilling other societal purposes, such as privacy.

In designing personal data-based AI, the following praxis is proposed as a basis of broad principles:

- Variations of concern over privacy among individual users should be an explicit part of data input points in algorithmic calculation. That is, an AI system must make inferences for affective benefits, such as trust and concern, so that algorithmic outputs

can be modulated according to different levels of concern and, ultimately, make effective use of data inputs.

- Systemic inferences and AI-based decisions must be confined within social contexts in which data are originally collected (i.e., maintaining their contextual integrity [Nissenbaum 2004, 2009] in the outcome of personal data analytics). Restricting algorithmic processing of data for any purpose other than the original contexts of data collection will reduce erroneous assumptions and enhance the efficiency of algorithmic outputs.
- Extraneous personal data points, such as race or socially sensitive status markers, should be excluded from algorithmic data analytics. This will not only limit the scope of data surveillance but also narrowly define the relevancy of data at stake, increasing the validity of AI-based decisions and inferences.
- In sum, the threshold of personal data input and output can be adjusted at the AI algorithm programming level based on social norms, value, and obligations, not purely on economic values and utility.

In this way, a pragmatic micro-level solution can emerge at the intersection of privacy norm, market, law, and design, as it can channel the data ecosystem into a normative goal. Thus, one attractive approach to policy is to take a more normative than technical approach, and a more consequential than instrumental view of personal data. As Napoli (2015) pointed out, algorithms can be regarded as institutions, like mass media, in that their effects are obligated to the public interest. Similarly, the protection of privacy in AI-based digital platforms should be thought of as an institutional product, whereby its interactions with users may carry public obligations under regulatory purview. On that basis, policymakers can justify regulatory intervention against an approach that relies on a utility-based, unregulated marketplace.

6.2. Norm Creation: Lessons from the Fight against Global Climate Change

This combination of all of these suggested measures won't be easy to put in place. The arguments against regulation of AI-based surveillance are many. A threat to the American value of freedom of expression is a ready one. In

addition, some observers view loss of privacy as the inevitable cost of creative innovation, adding that it is an individual's responsibility—not the government's or society's—to protect one's privacy. Skeptics point out that gridlock in Congress, or fierce industry capture and lobbying (which will intensify with the advancement of AI technologies), may prevent meaningful regulatory reform. Facebook employs dozens of internal and external lobbyists, with Alphabet Inc., Google's parent company, and Amazon employing even more. In 2017 Facebook spent roughly $11.5 million on lobbying in Washington, DC (Nix et al. 2018). What these facts show is the necessity of a cautious approach to policy on a comprehensive scale, not its abandonment. We must recognize that the resistance to regulatory instruments of privacy design, law, markets, and norms will seem insurmountable at the start.

The persistent argument of this book is that what is more evil than the evil itself is to sit out the contest and assume that good solutions will somehow win out through the magic of a free marketplace.

Here let's step back to see the essence of privacy issues, which become clearer when we compare them to issues like climate change. The readers of this book, despite their interest in surveillance and privacy, may not see them as comparable to the global climate crisis. Climate change is the man-made destruction of the ecology of nature and human beings, with irreversible damage. The policy debate is on how best to regulate carbon dioxide emission—the physical products of fossil fuels. Data, information flow, and AI algorithmic reasoning defy a traditional, industrial sense of crisis. In the United States, the public is divided over climate change, with a conservative fraction denying the scientific causes. Yet with respect to privacy, much of the US public remains concerned but perplexed, with no strong political stance about the causes and effects of surveillance.

The key differences between the two crises stop there, while similarities are many. First, both issues concern problems considered external to market solutions, as the respective industries cling to old habits—coal and surveillance. Second, just as climate change relates to almost every manufacturing sector across the globe, privacy is now a problem across all digital industries. Third, just as one's daily lifestyle and energy consumption pose a challenge to addressing climate change, a digital footprint is also correlated with consumption, as the core problem lies in the interaction between individuals and the industries that shape it. Fourth, the two issues mirror each other in terms of policy, given that other advanced nations

moved ahead with stricter regulatory standards than the United States, and the United States, particularly Silicon Valley, has been extremely resistant to the adoption of global standards, despite its status as one of the biggest perpetrators of both carbon emissions and privacy violations in the world. Sixth, there are shared cultural meanings, in that everyone is affected and no one can escape the negative consequences. If climate change is an existential threat to human beings, the threat to privacy is a comparable threat to the future of digital lives.

Why do these similarities matter? In the future, the desire for privacy will force us to ask how the scientific progress of technologies, and AI in particular, can be reconciled with social goals and demands. Just as global climate change has disturbed our physical ecosystem to an unforeseeable degree, our sense of identity, our ability to control data and define who we are—whether we are a "target" or a "waste" or a "good" or "bad" citizen—has been lost to an unpredictable degree, dismantled by a personal data-driven digital ecosystem. And AI is only worsening the trend.

Lessons from climate change show that meaningful changes to institutional surveillance practices and individual privacy behaviors require altering norms so that institutions, like Facebook, Google, and Amazon, and citizen-users, as well as policymakers and engineers, can internalize the value of privacy, just as we have seen substantial progress in regulatory practices related to climate change. Of course, this is a simplified version of the story. It took many years of international cooperation to reach the Paris Climate Accord in 2015 that installed a minimum-level restriction of carbon emission. Still, nations like the United States and Australia wanted to pull out of the agreement, with the United States withdrawing its participation in 2018. At the individual level, the resistance to behavioral and attitudinal change has been stiff; we now know how hard it is to persuade people to adopt a pro-environmental stance, that is, to recycle and conserve energy—not to mention getting them to accept scientific evidence.

Simply put, old habits die hard, individually or institutionally.

Privacy versus climate change in slowly boiling water

Any future regulatory models for privacy will be successful only with persistent support from each of the four parts (design, law, market, and norm) of the suggested measures, and this will be possible when a broad consensus about privacy, similar to that surrounding global climate change, is

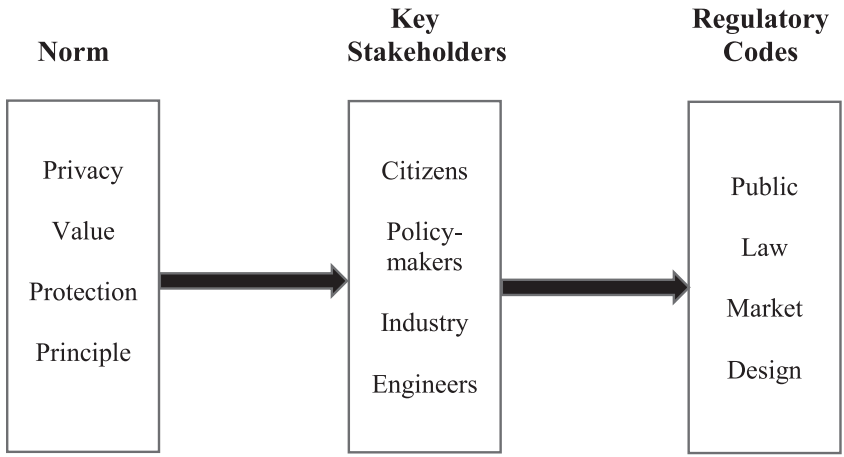

Figure 6.2. Sequential Steps toward Privacy Regulatory Codes

internalized by the industry, by citizen-consumer-users, and by policymakers alike.

Here norms refer, not only to public norms governing individual behavior in a narrow sense, but also to a comprehensive, if tacit, agreement among key stakeholders on a particular policy matter (e.g., Hamilton 2000, on media violence; Hamilton 2004, on broadcast news; Hasenclever, Mayer, and Rittberger 1997, on nuclear nonproliferation). Without consensus among key stakeholders, the marketplace, when left alone on its own, will treat privacy values, protection, and principles as externalities. Thus, the proposed policy alternatives must not be one-shot regulatory action imposed on top of industrial self-regulation. Instead, another step is needed before each constituent can develop a set of underlying principles and norms, and apply them as standard practices (Mueller 2009) (see Figure 6.2). The main contribution of EU GDPR in this regard, despite shortcomings discussed earlier, has been the creation of a broad global norm advocating a uniform level of privacy protection.

Al Gore, former vice president of the United States, in his warning against climate change, compared a human life to that of a frog who, though it could jump out, remains in water that is slowly being heated until it reaches a temperature that kills the frog. This analogy works for privacy as well. We may be sitting in a slowly boiling digital ecosystem, as surveillance over the last decades has intensified while surreptitiously

promoting reward, access, and convenience. A change of weather is noticeable in the physical world, but data surveillance and the loss of privacy are harder to detect. Thus, it may be difficult to internalize the negative consequences of data surveillance. If normative measures against a threat like climate change have been slow, difficult, resisted, and in some cases reversed, it is reasonable to expect that resistance to privacy protection—by both institutions and individuals—may be even worse.[4]

But the grand scale of our persistent problems demands an equivalent level of responses. Incremental regulatory reform will eventually invite a tide of fundamental changes.

CHAPTER 7

The Future of Digital Surveillance

> All the world is not, of course, a stage, but the crucial ways in which it isn't are not easy to specify.
>
> —Erving Goffman, 1967

7.1. Stories of Two Smart Cities—and a Yet Smarter One

A small city in Spain named Pontevedra, an old medieval town with a little more than 80,000 residents, had an innovative idea in 2010—to ban automobiles in the city center. This meant no need for parking or traffic signs. The benefit was enormous. Traffic accidents went to almost zero. The air quality improved drastically, and because people walked everywhere, tiny shops, restaurants, and bookstores began to flourish again; huge parking lots were replaced by common green areas. This was a citizen-centered model that was different from that of other municipal environments. Why does Pontevedra matter in our discussion of privacy? What does this vision of no technological intervention mean, as we ponder the future of digital surveillance?

Let's look at two others cities to see how shaping the urban environment can differ in essence. The first is Beijing. This Chinese megacity and the nation's capital announced an ambitious plan to monitor, record, and score individual citizen's behavior. The city government calls it a "social credit" system, with no sense of irony. People who commit misdeeds—street fights on trains, buses, or public places, for example—will receive low social credit scores and as a punishment be barred from certain public

transportation. Electronic eyes and facial recognition technology laid out all over the city will help the city government create a blacklist of problematic individuals, and this Big Brother list will also be available to third parties such as marketers, travel agencies, insurance brokers, and employers.

A North American version of this Orwellian smart city, this one in Canada, is not very different. This second example is Toronto, where Google's sister firm Sidewalk Labs planned to build on the once-abandoned waterfront.[1] Unlike the mechanism of restriction and punishment in Beijing, Toronto was going to use reward to obtain voluntary cooperation from residents under its AI-based surveillance. The urban planning in this collaborative project between Sidewalk Labs, and the Ontario government, with a budget of $50 million, would re-engineer the eastern side of Toronto by algorithms, tracking personal data to produce real-time, optimal municipal services. Various city-run AI applications, such as mobile parking apps, a traffic monitoring system, and hyperlocal weather sensors, would be calibrated to feed data to a database, which would be used to decide when to heat a snow-covered sidewalk—at the precise moment a pedestrian steps out—or to ease congestion in a public park according to the flow of automobile signals. By these benefits, Toronto would use infrastructure to improve the urban experience. Some said it would be a virtual city smarter than smart cities (Scola 2018), that is, algorithm-driven and feedback-rich—but with a repository of personal data that knows everything a pedestrian did on the boardwalk last night (Mosco 2015).

The idea here is not that there was a golden era of no surveillance technology to which we must return. Nor does this book advocate a Luddite technological phobia, claiming that we must destroy the monstrous artifacts promoting surveillance. The pedestrianization in Pontevedra was met with initial resistance. We can easily imagine complaints about the curtailment of individual freedom; others have pointed out that there are benefits to automobiles. Likewise, as much as there was resentment toward Sidewalk Labs, a private firm, taking over the electronic grid of Toronto, certain residents and local city officials welcomed economic revitalization and potentially better cash flow to the city. And it is not hard to see that some Beijing residents may like a "social credit" system for its social control, promotion of good citizenship, and prevention of petty street crimes.[2]

7.2. Social Construction of AI: Goffman's Public-Private Boundary Management

This book has come full circle to predict that the future of digital surveillance, as we consider the interaction between institutions and individuals (conditioned by a particular set of regulatory traditions in the United States), will manifest a resistance to a move toward privacy in the shaping of digital technologies, namely AI.

Since the business of predicting outcomes requires a multivariate model, I have relied on different units of analysis to reevaluate the impulses of both institutions and individuals. In this final chapter, I now turn to alternative versions of the possible future to better illuminate the current situation. It is clear that AI will be optimized for data surveillance, not privacy protection, as the appeal of no-nudge (i.e., regulatory nonintervention) lies in the seductive guile of free rational choice and self-determination. Given this momentum, one conundrum is that the forces driving surveillance only have reason to intensify, with no outside force to nudge the unregulated industry in a different direction. The contribution of this book is precisely to provide such a synthesis, which helps us ponder how to alter structural constraints on AI from the perspective of privacy protection.

Again, the central premise in this book is not about balance, nor about pros and cons, nor even about a cost-benefit analysis, but that the marketplace will not take care of human values and principles. It is true that there would be a short-term return on financial investment in a Google city. Tangible rewards would be evident. Employment in Toronto might increase, and smart-city services would entice citizens with their convenience; Beijing will have safer streets. Along with the best intended outcomes, however, comes a loss of privacy, of control over "selves" in human interaction, and of our abilities to define what is public and what is private about our lives.

Table 6.1 provides means to make this point. Compare City A, similar to Pontevedra, with City B, which could be Beijing or Toronto's Google city. The top-down surveillance structure in City B, unlike the bottom-up dynamics where citizens have a say, is coded to manufacture people's identities according to preprogrammed interests. The contextual flexibility in City A, where any face-to-face interaction gives subtle cues about the private and the public, does not close off social interaction algorithmically. In City B, humans are coded as data objects imposed by AIs, rather than

as subjects who can willfully condition the terms of their choices (Cohen 1999). A closed system like City B, where AI mediates all human activities, is hardly open to any interpretation of what is private or public, or to the freedom to determine who we are.[3]

It is painfully clear that the values and principles in the two sorts of city are different—analogue-driven Pontevedra, on the one side, and AI-augmented Beijing and Google city, on the other, where private moments in public spaces are no longer conceivable, banished by new technology in favor of surveillance. Reward and punishment will overwhelm an individual's ability to choose "being private." Erving Goffman's ideal of selective selves will be technologically impossible in Beijing and in a Google city. In shaping the direction of new technologies, however, we can learn from an alternative city constructed otherwise, like Pontevedra (Giddens 1983; see Neuman 2010; Pinch and Bijker 1984).

It is a collective choice about privacy's value and related norms and principles, and that choice will drive decisions about how we construct our world, its future, and how the combined determinants of economic and individual forces will shape digital ecosystems.

A fundamental is to put privacy and other values and interests, such as economics and efficiency, on a seesaw, as if a gain, let's say in privacy protection, would amount to a loss of the other value. It is easy to see why industry practitioners, executives, and policymakers in the US tradition of self-regulation would prefer "balance and trade-offs," because this zero-sum logic makes compromising privacy inevitable, as if determined by the inherent properties of digital technologies, and thus justifiable.

TABLE 7.1 Two Cities in Different Constructions

Characteristics	City System A	City System B
Main components	Human	AI machine
Public space	Analogue—open	Algorithm—closed
Private-public boundaries	Flexible	Fixed and manufactured
Administrative decisions	Bottom up	Top down
Data collection mechanism	Deliberate, social, and interactive	Punishment and reward
Context of social interaction	Situation-bound, face to face	AI mediated and context neglected
Citizens-users-individuals	Data subject	Data object
Model of privacy-surveillance	Goffman's life theater	Orwell's big brother

Source: Inspired by Neuman 1991; Van Dijk 2012.

Consequently, there would be no need to prioritize privacy over surveillance in product design—whether websites, smartphone apps, AI systems, algorithmic calculations, or smart cities. The key point is that surveillance results from deliberate decisions by policymakers and corporations, currently left unchecked, on whether people will be able to determine public-private boundaries (Nissenbaum 2004, 2009; Westin 1984).

7.3. The Business of Formulating Directional Hypotheses

Some say that it is foolish to try to predict the future, especially the future of digital technologies that may turn out to be different monsters than guessed. It is always tricky to claim knowledge about something that has not happened yet. Further, the anticipated direction of technology remains contingent on variables beyond the determinants of individuals and economics/political economy, whose interaction is further conditioned by political and regulatory forces. Nonetheless, the evidence pointing toward irreversible surveillance is overwhelming.

This book attempts to predict behavior at the aggregate societal level, while also directing readers toward the probable collective failure to stop the causes of surveillance. With respect to new technological development, an attempt to reverse surveillance will likely be unsuccessful, given (1) the inclinations among individual citizens-users, exacerbated by societal disparities, and (2) the pressures toward profit maximization and efficiency in new forms of digital platforms, advertising, and marketing. The central concern of this book has been that these underlying causal factors drive surveillance, that is, economics over individuals, as government policy sets the condition of their interplay. This book argues that, judging from the available evidence of web-based social interaction as well as emergent personal data-based AI systems, we are not likely to move toward privacy. In the foreseeable future of AI, we will fail, not only to craft the ideal of self-presentation, but also to rewire the norms that influence the choice of a technological design.

Privacy nihilism asserts that privacy is long gone; we have lost the battle for control; the fight for privacy is already over. Data brokers, retailers like Target, and credit companies like Equifax have collected, processed, and sold personal data ever since computer and electronic marketing began in the 1970s (Bogost 2018). This is long before Facebook, Google, Amazon,

and their AI algorithms. As Stuart Hall (2001) noted, the human gaze has always been a powerful tool of control, as it possesses the assertive power of surveillance in any human interactions; no complete sense of privacy, or the absence of others' surveillance, is possible in everyday life (Hall 2001; Westin 1998).

Although this book rejects a cynical pronouncement that privacy does not matter anymore, the individual citizen-user, despite occasional successes by those who oppose surveillance, has long lost control. This battle over privacy has been an inescapable human conditions since Adam and Eve (Acquisti et al. 2015; Miller 1971). We should qualify a position of nihilism, because the loss of control over privacy at the individual level does not preclude collective remedies, as proposed in chapter 6.

In a similar fashion, the thesis of surveillance capitalism is conditionally supported by the thesis of this book. The idea of surveillance capitalism—that personal data is being surveilled because it is a critical to revenue—encourages market-based control (Van Dijk 2014; see Zuboff 2019). The logic resonates with the fundamentals of market economics: just as the accumulation of capital is based on the exchange of physical goods/services, the collection and processing of personal data drives the emergent information-based economy in which all business activities and decisions hinge upon surveillance—that is, a largely tacit but effective power of big-data intelligence. The commodification of privacy in this regard is a natural growth of market capitalism, and extending this idea, we can understand how digital consumption, set up to be always correlated with surveillance, serves as a form of "free labor" performed by individuals (Gandy 1988; Park et al. 2018) that is largely invisible yet fuels personal information-based transactions. Google and YouTube search queries, for instance, are a kind of subsidy to Google AI in support of its intelligence and digital marketing, and thus of capitalism at large.

This is a powerful explanation of the economics of digital industries and their relationship to individual citizen-users. Putting aside the question of whether surveillance is genuinely a by-product of capitalism in an 18th-century, Adam Smithian sense, it gives an incomplete account of complex causal factors because it excludes a precise structural mechanism on the side of consumption. Oppressive surveillance is in fact an outcome of an intricate course of social psychological engagement that turns out to be a lot more complex than is often perceived. These dynamics were illustrated in chapter 3, which dissected passive and active privacy behavior,

which are intertwined and socially conditioned, as well as in section 5.3 of chapter 5, which revealed the convoluted psychological workings in AI-based environments.

People are submissive to data, but their action is voluntarily—and paradoxically—actively compliant because they cannot afford to be excluded from digital participation. They do give up privacy, not because they are captivated by the desire for hyperconsumption or because they are lazy or dumbfounded and do not care about privacy. Rather, people are vulnerable to the constant system of reward (e.g., free access) and punishment (e.g., no access) that limits their choices and actions (Foucault 1975), even among those with the most knowledge. Still, there is a significant line between individuals who are discreet, selective, and ready to deal with uncertainties within the constraint of digital ecosystem, and those who are reckless about divulging private details with no deliberate consideration. That line is socially incubated and is exacerbated by preexisting differences in societal positions, such as income and education—and, unfortunately, that is the basis of a pessimistic forecast about the future.

The title of this book could have been *Why and How We Will Lose the Battle over Personal Privacy*. In this sense, Karl Marx in *Capital* (1867) is partially right: the future outcome of any social change will be determined by the existing economic structure, but the techno-infrastructure in support of existing economic interest does not automatically lead to a monolithic mass of powerless individuals. Nor does it follow a fixed course of action. Ordinary people, though receptive, succumb to the convenience of inaction, but with occasional success at managing boundaries by turning off cookies and Wi-Fi, or refusing to provide sensitive personal data. In this sense, capitalism is a necessary but not sufficient condition for digital surveillance to dominate the fight for privacy. By focusing on complex individual determinants, we can gain insight into how the social psychological underpinnings of individuals become endogenous to the force of surveillance. It is a subtle process. And it may be counterintuitive to identify complicated, intertwined causalities when the net outcome is obvious. But to fix the problem, we must understand why the outcome is obvious. There is no reason to be optimistic about the future course of surveillance when no influence is exerted on the powerful determinants we have analyzed.

7.4. The Decisive Epoch: Look Back to Look Ahead

A central thesis of this book can be epitomized in the past development of one technological artifact, that is, HTTM cookies, which are small electronic data files used to track web browsing. As historians insist, the best clue about the future is in the past.

It was June 1994. A computer programmer named Lou Montulli at Netscape Communications developed the authentication code. The idea was to verify user ID so that a website could know if the person visiting it had already visited (Schwartz, J. 2001). Cookies were an effective way to obtain this knowledge—a reliable way to send and store unique identifying information in each user's computer. It was like a caller ID, but in this case, the relationship between a caller and a receiver was not symmetrical, because electronic communication occurs one way, with cookies flowing from the website/company to the user/person who is to be identified.[4]

It would be oversimplification to blame cookies, as much has changed since they were created. But this banal decision in 1994 turned out to be decisive to the power of digital surveillance. One can imagine that a Silicon Valley corporation might assign a data scientist to code its AI to remember forever the digital footprint of every single user. What would the scientist, one is privacy conscious and concerned, do? What if the scientist had an opportunity to stand up against the corporate decision to create a code to surveil? What if the scientist's answer to the corporation's bottom-line pressure was no?

Such a scientist might understand the cognitive burden imposed on individuals in their digital consumption because no one reads the legally convoluted fine print of the privacy statements that can be found several clicks away from the front pages of most websites (Park 2011a, 2011b; Turow 2003, 2005). A user will find the terms of data use and services ambiguous or upsetting. No wonder, more than 20 years after the first cookies started to track us, a majority of the US consumers still do not know what they are (Madden et al. 2015). Now there are third-party cookies, persistent cookies, supercookies, and so on. No law, with regard to their regulation, was passed. Nor did the social psychology of ordinary people change. Nor did the corporate impulse to track down the minute clickstreams of individual consumers. Still, what if? What would happen if the data scientist reversed the code, in favor of privacy? Can we imagine such

a possibility? Or should we remain hopeless because natural inclinations appear too great to resist?

Cookies, algorithm, and AI: Norms and principles for deliberate choices

The course of surveillance history could have been otherwise. Think of cookies in Web 1.0 as a precursor of personal data-based AI, or if the augmented version of cookies in Web 3.0, or even of the Fourth Industrial Revolution. Thomas Hughes (1987) referred to the notion of a reverse salient, a technological element or sticking design problem holds back the progress of an entire system. We can perceive cookies as a system element, designed to outstrip the capacity of human beings to control their personal privacy in favor of the greedy corporate instinct to surveil. In essence, a banal smart-code decision at the very birth of the digital ecosystem set the tone for the future. By now, to recode and rewire the entire system is daunting. But it was a quarter century ago. HTTM cookies sparked FTC hearings in 1996 and 1997, but the discussion got buried by Netscape, just as 2018 congressional hearings about Facebook–Cambridge Analytica and its role in suppressing voter turnout fell into oblivion amid sensational media coverage about Mark Zuckerberg and his performance in front of US senators.

Cookies were a deliberate choice, as much as AI is now. This choice can be reversed to the extent that one element of a system be more salient than another, as seen in Pontevedra's banning automobiles. Such a deliberate outcome can be constructed if sufficient political will and collective norms precede intervention via policy. And this possibility, as this book argues, will be clearer to us when we realize there is no inherent reason for future digital technologies to operate in a particular way.

Moments ripe for pushing reform are rare, but do occur. And when they do, their impact is profound. In April 2018, such a moment occurred when thousands of Google workers signed a letter to protest the company's role in Project Maven, a Pentagon-led drone project seeking to develop better human targeting based on Google Cloud's recognition AI software. Several months later, yet another Google subsidiary's bidding on the Joint Enterprise Defense Infrastructure (JEDI), a massive military computing contract, was dropped because employees expressed grave concern that the project could violate Google's own AI principles. On this occasion,

employees at Microsoft also walked out in protest against the company's involvement in the same project.

The tech employees' urge for change may be short-lived, an episode. Yet these moments hint at the possibility of reorganizing principles, creating values that could be internalized in a Silicon Valley–dominated tech sector. The change may not be fast enough, as surveillance is natural in the marketplace. Accordingly, the pace of history will probably move in less than ideal fashion. Nevertheless, we can agree on the fundamental norms, on the need for reshaping surveillance technologies, and on how to achieve the change.

I began this book by considering the NSA PRISM surveillance made public by Edward Snowden in 2013. This concentration of digital surveillance power started with Netscape and FTC decisions on cookies in 1996 and 1997, which produced no regulatory discipline. Nowadays Amazon's cloud, for instance, is so much bigger than that of the US government that the latter must rely on it. And the blurred relationship in intelligence surveillance between the private and public sectors came to critical light because of Edward Snowden, a private citizen, a former CIA contractor, and an individual who changed the course of surveillance.

History repeats, but only under certain conditions. Voluntary, conscious, and organizing efforts by those on the front line in Silicon Valley and in other technology workforces might occur, and we might see familiar patterns of institutional practices begin to break down. Snowden did not articulate his ideal in terms of Goffman's self-presentation, but Snowden's beliefs seem opposed to life under the alternative, an Orwellian fear. Imagine, for a moment, a future in which a few individuals, like Edward Snowden, translate their beliefs on privacy into collective wisdom and action.

Appendix A

The Locus of Privacy Protection in the Marketplace

The main findings reported in chapter 2, section 2.2, are based on a combination of top and random samples. The sample pool was created as follows. The US sites were first identified from the top 500 global websites. Then 500 websites were randomly selected from the first 10,000 AOL search queries.

The AOL log ensures the variance of sampled sites in the externally valid internet universe, while the inclusion of the top sites incorporates the most visited venues as in a daily context. Thus, the sample pool of 1,000 websites were created with top and random samples (500 + 500) combined. Duplicates between the top and the random sites were excluded from the pool. From the sample of top websites, one government-operated site and three sites with the same policies were excluded. A site with a U.S. Internet Protocol address, but operating under foreign ownership, was eliminated, creating a total of 148 sites for analysis. For the random sample (n = 250), the following multistage cluster sampling was used. In the first stage, 500 clusters of individual search queries were identified by randomly selecting them from the 10,000 AOL user batches. In the second stage, an individual URL within each cluster was randomly selected. Each cluster was mutually exclusive, consisting of 20–70 unique URLs. With a total of 500 clusters, this selection includes 10,000–35,000 sites from which to select the final samples.

Note the advantage of this technique in increasing the chance of equal selection when it is impossible to locate all the elements within the sample

frame. Three broken URLs were identified and eliminated during the coding process. The sampling rate was (0.25) * (0.01), with the confidence level of 95 percent and a SE ±4.9. It is important to note that the FTC in its own study (1998) applied a similar technique using a random sample, with the author of this book following the logic of the FTC study for consistency.

There were the two levels in the explanatory variables: (1) the market domain, and (2) the site factors (Table A.1). The market factors measured whether the sites in a specific market domain are in fact more inclined to provide privacy protection functions. The site factors measured the influences of individual website attributes, such as financial resources and the number of years in site operation, in incentivizing further provision. The study also examined whether better-resourced sites, as indicated in revenues, traffic ranking, seal membership, or broader business scopes, are more responsive to the demand from the public in the marketplace.

Key measures of privacy concern, as reported in Figure 3.6 (1990, 2000, 2003, 2006, 2008, 2014, and 2015), are based on a composite of several sources:

TABLE A.1 Sample Characteristics

Factors	Mean	Std. Dev.	Descriptions
Market Factors			Type of market domain
Online	.63	.48	Whether the site operation is confined online (1 = yes, 0 = no)
New Media	.10	.29	Search engine or directory sites (1 = yes, 0 = no)
Sensitive 1	.08	.26	Whether a site deals with sensitive data (health or financial information) (1 = yes, 0 = no)
Sensitive 2	.08	.27	Whether a site is targeted toward children, teenagers, or younger users (1 = yes, 0 = no)
Site factors			Characteristics of an individual site
Publicly listed	.30	.46	The site (or its parent company) in public stock market (1 = yes, 0 = no)
Ranking	56978.3	302953.7	Traffic ranking in September 2008 (000,000)
Year (new)	10.24	3.50	Number of years of operation
US percent	62.38	26.38	Percent of US users (%)

Note: Data collection was made in 2008.

- 1990 Harris Poll, January 1990 (Harris study no. 892049)
- USA Weekend poll, *USA Weekend* Magazine, July 2, 2000
- 2003 Harris Poll, March 19, 2003
- USA Today / Gallup Poll, Government Phone Records Reaction, *USA Today*, May 2006
- Pew Internet and American Life Project: Cloud Computing Raises Privacy Concerns, September 12, 2008
- US attitudes toward the "Right to be Forgotten," IndustryView 2014, Software Advice, September 5, 2014
- Pew Internet and American Life Project: Americans' Attitudes about Privacy, Security and Surveillance, May 20, 2015

Primary sources for privacy protection in Figure 3.5 are the following:

- FTC, Privacy Online: A Report to Congress (1998).
- FTC, Self-Regulation and Privacy Online, before the House Commerce Subcommittee on Telecommunications, Trade, and Consumer Protection (1999)
- FTC, Privacy Online: Fair Information Practices in the Electronic Marketplace: A Report to Congress (2000)
- FTC, Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers (2010)
- LaRose and Rifon 2006, data collection done in 2002

The source for Privacy Protection among Tech Companies in Table 3.2 is the following:

- New America Foundation 2017

Specific methodologies in these reports can be found at https://thedigitalstandard.org.

Appendix B

Empirical Evidence in Two Strands

The concepts, variables, and findings reported in chapter 3, section 3.2, are primarily based on an analysis of the Knowledge Networks panel (2008). The Knowledge Networks panel was based on a national probability sample of adult internet users (18 and older) who were recruited using random digit dialing on a sample of all US residential telephone numbers. The panel participants were asked to visit a site and complete an online survey. In order to ensure representativeness, the sample was drawn to reflect demographic distributions in key census areas: income, gender, age, and region. Demographic characteristics of the sample are similar to those of the general population reported in the US Census Bureau's 2007 American Community Survey (ACS). With respect to educational attainment, the median education level for those 25 or older in both data sets is some college. Household income (the medians in the ACS and the current study are $50,000–74,999 and $60,000–74,999, respectively), gender (females in the ACS and the sample are 52.4% and 53.6%, respectively), and age (the median age for those 18 or older in the ASC and the current study is 45–54 and 47, respectively) resemble the profiles of the general population.

B.1. Knowledge Networks Study: All Measures Analyzed

Knowledge/literacy measures used in chapter 3, section 3.2, are in Table B.1. These items were based on a factual knowledge-test, constructed on

a binary scale (correct = 1; incorrect and DK = 0), instead of observing self-assessed cognitive efficacy. The knowledge measures were used to create knowledge index scores in two areas: policy understanding and surveillance awareness, representing distinctive but related cognitive dimensions. Dependent variables of privacy control were elaborated into social and technical dimensions. This follows multifaceted nature of information control, which requires a combination of social and technical skills as intertwined in internet uses. On a six-point scale, respondents reported the extent to which they were involved in information control, ranging from never to very often. Eight items were used for the social dimension, and the study measured four items for the technical dimension, modified from the extant literature (Acquisti and Grosslags 2005; Marx 1998; Metzger 2007; Turow 2003, 2005; Turow et al. 2012). Those are Avoidance (M =3.21, SD = 1.85), Hiding (M = 2.54, SD = 1.73), Withdrawal 1 (M = 3.42, SD = 1.72), Withdrawal 2 (M = 4.28, SD = 1.63), Complain (M = 1.50, SD = 1.07), Rectify 1 (M = 3.51, SD = 1.82), Rectify 2 (M = 3.58, SD = 1.97), and Multiple account (M = 2.89, SD = 1.97) in the social dimension; Clearing history (M = 3.49, SD = 1.81), Filtering email (M = 4.56, SD = 1.90), Erasing cookies (M = 3.68, SD = 1.90), and Using privacy protection software (M =1.41, SD = 1.48) in the technical dimension.

In Figure 2.2, informational privacy concern, as opposed to institution-related privacy concern (e.g., specific sectors or entities, such as Google vs. Apple), was the measure of a person's concern over different aspects of information surveillance: (1) data collection (M = 2.74, SD = 1.35) and (2) data appropriation (M = 3.57, SD = 1.20). Users were asked, on a six-point scale, about their agreement with a statement assessing anxiety level, anchored by strongly agree and strongly disagree. The items were combined into one measure to capture worry, and this was a result of factor analysis with a principal component matrix of .80 (eigenvalue of 1.92, explained variance of 48.07%, Cronbach alpha = .63). Finally, the measure of willingness for trade-off asked the level of likelihood to trade off different types of personal data for financial gain or access to favorable content, on a six-point scale (anchored by not at all likely and very likely). Then, each item was listed to rate the likelihood on a six-point scale. Nine data items were drawn from Ackerman, Cranor, and Reagle 1999; Acquisti and Grosslags 2005; Culnan and Armstrong 1999; and Ribak and Turow 2003. An additive index (M = 20.31, SD = 9.52, Cronbach alpha = .86) was created based on a composite scores of these nine items to capture the

TABLE B.1 Privacy Knowledge / Literacy Index Scores

M	SD	Items
Surveillance awareness		
0.75	0.43	Companies today have the ability to place an online advertisement that targets you based on information collected on your web-browsing behavior.
0.57	0.49	A company can tell that you have opened an email even if you do not respond.
0.65	0.47	When you go to a website, it can collect information about you even if you do not register.[a]
0.66	0.47	Popular search engine sites, such as Google, track the sites you come from and go to.
0.45	0.49	E-commerce sites, such as Amazon or Netflix, may exchange your personal information with law enforcement and credit bureau.
0.72	0.44	What a computer user clicks while online surfing can be recorded as a trail.
0.68	0.46	Most online merchants monitor and record your browsing in their sites.
0.25	0.43	When a website has a privacy policy, it means the site will not share your information with other websites or companies.
Policy understanding		
0.20	0.40	Government policy restricts how long websites can keep the information they gather about you.
0.22	0.41	It is legal for an online store to charge different people different prices at the same time of day.
0.40	0.49	A website is legally allowed to share information about you with affiliates without telling you the names of the affiliates.
0.14	0.35	By law, e-commerce sites, such as Amazon, are required to give you the opportunity to see the information they gather about you.[b]
0.20	0.40	Privacy laws require website policies to have easy-to-understand rules and the same format.
0.56	0.49	US government agencies can collect information about you online without your knowledge and consent.
0.22	0.41	When I give personal information to an online banking site such as citibank.com, privacy laws say the site has no right to share that information, even with companies it owns.

Note. Analyzed in Figure 2.5 is

[a] Surveillance Awareness Transfer, analyzed in Figure 2.5.

[b] Policy Understanding Appropriation, analyzed in Figure 2.5.

extent to which users perceive the likelihood of divulging data for reward or benefit at hand.

B.2. Privacy Mobile Study: Mobile Privacy Knowledge

For the mobile sample, a series of in-depth interviews were conducted along with survey analysis. A composite data set was constructed after the response validity check in the first wave of data collection, adding a new set of data from the second wave in a series of pilot studies throughout 2011–2012. The analyses were based on 60 individual observation sessions. The study population was recruited using nonprobability purposive and snowballing sampling procedures to seek young African American adult users / underrepresented communities. Snowballing sampling has advantages in recruiting members of underprivileged communities (Burrell 2010). In addition, because young users often rely on a network of knowledgeable associates or peers, this technique provided us with an effective reference point to understand a target group. Descriptive characteristics of the participants are as follows: parental education (M = 3.36, SD = 1.13 on a six-point scale); age (M = 20.23, SD = 2.07); household income (M = 4.19, SD = 1.53 on a six-point scale); and gender (high: female, 53.3%).

Mobile knowledge was operationalized as user awareness in the two dimensions of institutional practices: (1) information and (2) location-related mobile personal data. For both dimensions, the participants were asked seven true-false questions that rated their understanding of mobile-based surveillance practices. Items were adopted from prior studies (Madden and Smith 2010; Turow 2005) and were later coded 1 for correct answers with 0 assigned to all other responses.

See Table B.2 for the distribution of individual items (KR 20 = .57). One possible concern is that the data sets need further updates, especially given recent changes in digital environments. But again, two things need to be kept in mind: first, the demographic characteristics of the US population have remained more or less the same since 2008; second, in principle, behavioral and cognitive change tends to occur extremely slowly. Even when we account for new technologies, there seems no apparent reason (or documented evidence) for the mass public to change privacy behavior. Still, future studies should update the findings reported in this book with new knowledge and behavioral measures reflective of algorithm- and AI-based digital devices.

TABLE B.2: Mobile Privacy Knowledge / Literacy Index Scores

M	SD	Items
Information related		
0.84	0.36	Most mobile apps, such as Facebook or Yahoo, monitor and record your browsing.
0.66	0.47	Companies today have the ability to place an ad that targets you based on information collected on your mobile phone.
0.42	0.49	When a mobile app has a privacy policy, it means the app will not share your information with other companies.
0.33	0.47	A mobile app service is legally allowed to share information about you with affiliates without telling you the names of the affiliates.
0.18	0.39	Government policy restricts how long mobile or smartphone service providers, such as Google Phone or iPhone, can store your personal data.
Location related		
0.52	0.50	Carrying cell phones gives law authorities the ability to track the places you go.
0.57	0.49	It is legal for your mobile or smartphone service provider, such as Apple (iPhone), to collect your location when you use your mobile phone.

Notes

Chapter 1

1. In describing the technological change, I use the term *digitalization* in this book. However, one could simply use *digitization*, which denotes the process of converting analog or physical objects into a digital format. In general, I use the terms interchangeably, but in the discussion of AI in chapter 5, the technical use of *digitization* is more proper.

2. This organizational logic follows the book's analytical aims. The first two chapters in Part II raise the foundational issues (of people and corporate institutions, respectively, as conditioned by policy) and then apply the insights derived there to algorithm-based AI in chapter 5, followed by policy suggestions in chapter 6. Finally, chapter 7 synthesizes what has preceded in the book's technological forecasts.

Chapter 3

1. *Social psychology* in this book is used broadly. On one hand, there is a sociological tradition that helps us understand social stratification and disparities of knowledge, skills, and behavior (DiMaggio et al. 2001; Hargittai, 2002, 2006). On the other hand, I also tap into user psychology as dissected through convenience, inaction, and passively succumbing to the enticement of reward. This combination is synthesized later, in chapter 5, to reveal why users' resistance is most likely to fail amid their consumption of AI-created options.

2. Rationality remains a contentious issue. especially because US policy as well as private firms assume a perfect user knowledge. The notion of rationality concerns individual cognition (behavioral psychology), but the fact that the acquisition of knowledge and skill sets is socially incubated is a significant sociological insight guiding privacy studies.

Chapter 4

1. Laissez-faire, in a strict sense, is not the same as self-regulation, since it implies the lack of regulation. In fact, data privacy and surveillance issues are subject to limited sector-level regulations like HIPAA (the Health Insurance Portability and Accountability Act, 1996), the Fair Credit Reporting Act (1970), and the Gramm-Leach-Bliley Act (1999), particularly in the financial and medical sectors. Setting aside the questionable applicability of those laws to AI and digital advertising, the precise choice of wording is not of much significance because the essential concern involves a largely hands-off approach that has resulted in a market failure regarding privacy protection—like the parallel market failure regarding climate change.
2. It should be noted that consumer protection and market competition have been the two main goals of policy regulation not only in the realm of privacy protection, but also on a range of other issues such as telecommunication infrastructure and mass-media content.

Chapter 5

1. The terms *AI* and *algorithm* are used here as if interchangeable, though they are not precisely the same. Algorithms can be considered calculative software that is often used to fuel data surveillance, whereas AI is much like a machine that produces decisions based on algorithmic reasoning or a programmed decision process. The distinction is an important one, but I reserve extensive debate for future work. See related details in Hall and Pesenti 2017.

Chapter 6

1. Lawrence Lessig, in the original version of his book *Code*, argued for a property-based solution on privacy. I reject his fundamental rationale here (see also Rotenberg 2001), though I adopt some of Lessig's general propositions about code.
2. It is important to describe the AI industry's relationship with firms in the new-media industry. Not every new-media company is based on AI, but AI has become an integral, if not the most significant, part of such firms as Google, Netflix, and Facebook. In some cases, a certain AI technical capacity may be independently developed but then sold to and merged into bigger companies, while in-house AI at the larger companies (Google X, for instance) is being developed at the same time. Thus, this distinction between the AI industry and the new media industry, though important, becomes blurred in reality. One key concern presented in this book, the possibility of AI integration creating wide-ranging personal data platforms, is directly connected to this agglomeration of AI capacity within a handful of companies.
3. In fact, the second and third proposals have been put forth, but to no avail. The reason for their failure as policy options is not entirely clear and may be attributable to a multitude of confounding factors. However, as I argue in this book, the fundamental constraints of institutional and individual impulses have been inherently inhibiting.

Regardless of bureaucratic hurdles to regulatory change at either the design or the implementation stage, a collective will to achieve policy reform must congeal among policy activists to generate sufficient pressure for change.

4. Fuchs (2015) in his critical work wisely suggested the implementation of an alternative noncommercial model, such as Wikipedia, in which there exists no pressure to attract advertisers. This is a smart suggestion that shares the fundamental concern in this book regarding financial incentives—coupled with AI transition—that motivate personal data surveillance. The difference between his suggestion and this book's proposal, however, is that here I emphasize the regulatory norms that must precede such a radical transformation.

Chapter 7

1. Public backlash against AI surveillance led Toronto and Sidewalk Labs to reconsider the Google city project as of August 2019 (Wong 2019). The Google project had begun to be scaled down, when Sidewalk Labs eventually changed much of its initial project plan in 2020, citing the Covid-19 pandemic and related complications. However, the fundamental point raised here remains relevant, in that more urban projects are expected in future with data surveillance as a core function of smart cities. In the U.S. alone, municipal governments already implemented a number of similar projects, such as a smart parking system in San Francisco and Philadelphia's smart city initiative, which incorporates the Automatic Vehicle Location (AVL) system collecting location data from buses to prevent disrupting commuter schedules.

2. This book has mostly focused on US institutions, policies, and AI-based digital platforms that grew out of Silicon Valley, but considering examples outside the United States could yield new insights. Policy studies have widely adopted this comparative approach. Here the purpose of comparing smart cities in Italy, China, and Canada is to illuminate the controversies about US-based Google and to show that the issue is not unique to the United States. Moreover, the future AI-based transition is likely to look much like one of the choices made by each city, in which the extent of data surveillance is consciously designed.

3. Roughly put, the two models can be understood as the difference between "opt in" and "opt out" at the user-interface level. This is a useful distinction in which the latter option assumes omnipresent surveillance that entraps everyone, whereas the other system allows such data collection only when a consumer specifically indicates a desire for it. No matter how crude it might sound, some variant of an opt-in model, especially with gradient data input points (again, depending on the specific contexts a person designates) in an AI-driven digital ecosystem, will take a step toward Goffman's ideal and might germinate alternative business models.

4. To be precise, cookies are stored on a user's personal computer, and the web browser sends the cookie back to the server, thus making this theoretically a two-way system. However, most users do not even know that this is occurring, and even when they do, the flow of data back to the server happens regardless of their intention or communicative goals. I am willing to entertain alternative interpretations, however.

References

Abdullah, S., and Choudhury, T. (2018). Sensing technologies for monitoring serious mental illnesses. *IEEE MultiMedia 25*(1), 61–75. https://doi.org/10.1109/MMUL.2018.011921236

Ackerman, M., Lorrie, C., and Reagle, J. (1999). Privacy in e-commerce: Examining user scenarios and privacy preferences. *Proceedings of the ACM Conference in Electronic Commerce*, November, 1–8.

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the ACM electronic commerce conference* (EC04), 21–29.

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science 347*(6221), 509–514. https://doi.org/10.1126/science.aaa1465

Acquisti, A., and Grosslags, J. (2005). Privacy and rationality in decision making. *IEEE Security and Privacy 3*, 26–33. https://doi.org/10.1109/MSP.2005.22

Agre, P. (1998). *Technology and privacy: The new landscape*. Cambridge, MA: MIT Press.

Altman, I., Vinsel, A., and Brown, B. B. (1981). Dialectic conceptions in social psychology: An application to social penetration and privacy regulation. In *Advances in experimental social psychology*, 14, 107–160. Cambridge, MA: Academic Press.

Ananny, M., and Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society 20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Anderson, C. (2004). The long tail. *Wired Magazine 12*(10), 170–177.

Andrade, E. B., Kaltcheva, V., and Weitz, B. (2002). Self-disclosure on the web: The impact of privacy policy, reward, and company reputation. In S. M. Broniarczyk and K. Nakamoto (eds.), *ACR North American Advances*, 29, 350–353. Valdosta, GA: Association for Consumer Research Pages.

Anonymous. (1998). To reveal or not to reveal: A theoretical model of anonymous communication. *Communication Theory 8*(4), 381–407. https://doi.org/10.1111/j.1468-2885.1998.tb00226.x

Awad, N. F., and Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quarterly 30*(1), 13–28. https://doi.org/10.2307/25148715

Ball-Rokeach, S. J., and DeFleur, M. L. (1976). A dependency model of mass-media effects. *Communication Research 3*(1), 3–21. https://doi.org/10.1177/009365027600300101

Barocas, S., and Selbst, A. D. (2016). Big data's disparate impact. *California Law Review 104*, 671. https://doi.org/10.2139/ssrn.2477899

Barton, G., Resnick, P., and Turner Lee, N. (2018). Algorithmic bias detection and mitigation: Developing industry best practices. Panel presentation, TPRC, September 21. Retrieved from https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

Baruh, L., Secinti, E., and Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication 67*(1), 26–53. https://doi.org/10.1111/jcom.12276

Bates, B. J. (1993). Concentration in local television markets. *Journal of Media Economics 6*(3), 3–21. https://doi.org/10.1080/08997769309358242

Bazarova, N. N. (2012). Public intimacy: Disclosure interpretation and social judgments on Facebook. *Journal of Communication 62*(5), 815–832. https://doi.org/10.1111/j.1460-2466.2012.01664.x

Bellotti, V., and Samarajiva, R. (1998). Interactivity as though privacy mattered. In P. Agre and M. Rotenberg (eds.), *Technology and privacy: The new landscape*, 277–310. Cambridge, MA: MIT Press.

Benkler, Y., Faris, R., and Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford: Oxford University Press.

Beniger, J. (2009). *The control revolution: Technological and economic origins of the information society*. Cambridge, MA: Harvard University Press.

Bennett, C. J., and Raab, C. D. (2018). Revisiting the governance of privacy: Contemporary policy instruments in global perspective. *Regulation & Governance 27* (September). https://doi.org/10.1111/rego.12222

Bogost, I. (2018). Welcome to the age of privacy nihilism. *The Atlantic*, August 23. Retrieved from https://www.theatlantic.com/technology/archive/2018/08/the-age-of-privacy-nihilism-is-here/568198/

Bollier, D., and Firestone, C. M. (2010). The promise and peril of big data. Washington, DC: Aspen Institute, Communications and Society Program. Retrieved from https://assets.aspeninstitute.org/content/uploads/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. London: Routledge.

Bourdieu, P. (1990). *The logic of practice*. Palo Alto, CA: Stanford University Press.

Bowker, G., and Star, S. L. (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

boyd, d. (2018). What hath we wrought? SXSW EDU. Retrieved from https://www

.sxswedu.com/news/2018/watch-danah-boyd-keynote-what-hath-we-wrought-video/

boyd, d., and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society 15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

boyd, d., and Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday 15*(8). Retrieved from https://firstmonday.org/article/view/3086/2589

boyd, d., Hargittai, E., Schultz, J., and Palfrey, J. (2011). Why parents help their children lie to Facebook about age: Unintended consequences of the "Children's Online Privacy Protection Act." *First Monday 16*(11). Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/3850

Brandimarte, L., Acquisti, A., and Loewenstein, G. (2012). Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*. Retrieved from http://spp.sagepub.com/content/early/2012/08/08/1948550612455931.

Büchi, M., Just, N., and Latzer, M. (2016). Enhancing online privacy at the user level: The role of internet skills and policy implications. AoIR Selected Papers of Internet Research, 6. Retrieved from https://www.researchgate.net/publication/317348389_Enhancing_online_privacy_at_the_user_level_The_role_of_Internet_skills_and_policy_implications

Budak, J., and Rajh, E. (2018). Citizens' online surveillance concerns in Croatia. *Surveillance & Society 16*(3), 347–361. https://doi.org/10.24908/ss.v16i3.6907

Burke, K. (1989). *On symbols and society*. Chicago: University of Chicago Press.

Cadwalladr, C. (2018). "I made Steve Bannon's psychological warfare tool": Meet the data war whistleblower. *The Guardian*, March 18. Retrieved from https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-facebook-nix-bannon-trump

Cairncross, F. (1997). *The death of distance: How the communications revolution will change our lives*. Cambridge, MA: Harvard Business Review Press.

Campbell, S. W., and Park, Y. J. (2008). Social implications of mobile telephony: The rise of personal communication society. *Sociology Compass 2*(2), 371–387. https://doi.org/10.1111/j.1751-9020.2007.00080.x

Castells, M. (1997). *Power of identity*. London: Blackwell.

Castells, M. (2002). *The internet galaxy: Reflections on the internet, business, and society*. Oxford: Oxford University Press.

Caughlin, J. P., and Petronio, S. (2004). Privacy in families. In Anita L. Vangelisti (ed.), *Handbook of family communication*, 379–412. Mahwah, NJ: Lawrence Erlbaum Associates.

Chen, H. T. (2018). Revisiting the privacy paradox on social media with an extended privacy calculus model: The effect of privacy concerns, privacy self-efficacy, and social capital on privacy management. *American Behavioral Scientist 62*(10), 1392–1412. https://doi.org/10.1177/0002764218792691

Chen, W., Huang, G., Miller, J., Lee, K. H., Mauro, D., Stephens, B., and Li, X. (2018). "As we grow, it will become a priority": American mobile start-ups' privacy practices. *American Behavioral Scientist 62*(10), 1338–1355. https://doi.org/10.1177/0002764218787867

Chen, W., Quan-Haase, A., and Park, Y. J. (2018). Privacy and data management: The

user and producer perspectives. *American Behavioral Scientist 62*(10). https://doi.org/10.1177/0002764218791287

Chen, W., and Wellman, B. (2005). Minding the cyber-gap: The internet and social inequality. In *The Blackwell Companion to Social Inequalities*, 523–545. https://doi.org/10.1002/9780470996973.ch23

Cheney-Lippold, J. (2017). *We are data: Algorithms and the making of our digital selves*. New York: NYU Press.

Cho, H., Lee, J. S., and Chung, S. (2010). Optimistic bias about online privacy risks: Testing the moderating effects of perceived controllability and prior experience. *Computers in Human Behavior 26*(5), 987–995. https://doi.org/10.1016/j.chb.2010.02.012

Cohen, J. E. (1999). Examined lives: Informational privacy and the subject as object. *Stanford Law Review 52*, 1373–1438. Retrieved from https://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1819&context=facpub

Cohen, J. E. (2012a). What privacy is for. *Harvard Law Review 126*, 1904–1933. Retrieved from https://cdn.harvardlawreview.org/wp-content/uploads/pdfs/vol126_cohen.pdf

Cohen, J. E. (2012b). Irrational privacy. *Journal on Telecommunication & High Technology Law Review 10*, 241. Retrieved from http://www.juliecohen.com/attachments/File/CohenWhatPrivacyIsFor.pdf

Couldry, N., and Mejias, U. A. (2018). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media 20*(4), 336–349. https://doi.org/10.1177/1527476418796632

Couldry, N., and Turow, J. (2014). Advertising, big data and the clearance of the public realm: Marketers' new approaches to the content subsidy. *International Journal of Communication 8*, 1710–1726. Retrieved from https://repository.upenn.edu/asc_papers/413

Crawford, K. (2012). Four ways of listening with an iPhone: From sound and network listening to biometric data and geolocative tracking. In L. Hjorth, J. Burgess, and I. Richardson (eds.), *Studying mobile media*, 221–236. London: Routledge.

Crawford, K. (2013). The hidden biases in big data. *Harvard Business Review 1*. Retrieved from https://hbr.org/2013/04/the-hidden-biases-in-big-data

Crawford, S. P. (2013). *Captive audience: The telecom industry and monopoly power in the new gilded age*. New Haven, CT: Yale University Press.

Culnan, M. J., and Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science 10*(1), 104–115. https://doi.org/10.1287/orsc.10.1.104

Czitrom, D. (1982). *Media and the American minds: From Morse to McLuhan*. Chapel Hill: University of North Carolina Press.

Dahlgren, P. (2001). The public sphere and the net: Structure, space, and communication. In L. Bennett and R. Entman (eds.), *Mediated politics: Communication in the future of democracy*, 33–55. Cambridge: Cambridge University Press.

Danna, A., and Gandy, O. H. (2002). All that glitters is not gold: Digging beneath the surface of data mining. *Journal of Business Ethics 40*(4), 373–386. https://doi.org/10.1023/A:1020845814009

Delli Carpini, M. X., and Keeter, S. (1996). *What Americans know about politics and why it matters.* New Haven, CT: Yale University Press.

Dienlin, T., and Metzger, M. J. (2016). An extended privacy calculus model for SNSs: Analyzing self-disclosure and self-withdrawal in a representative US sample. *Journal of Computer-Mediated Communication 21*(5), 368–383. https://doi.org/10.1111/jcc4.12163

DiMaggio, P., Hargittai, E., Neuman, W. R., and Robinson, J. P. (2001). Social implications of the internet. *Annual Review of Sociology 27*(1), 307–336. https://doi.org/10.1146/annurev.soc.27.1.307

Dutton, W. H., and Blank, G. (2015). *Cultures on the Internet 42*(4–5), 55–57. Retrieved from https://ssrn.com/abstract=2545596

Dutton, W. H., Guerra, G. A., Zizzo, D. J., and Peltu, M. (2005). The cyber trust tension in E government: Balancing identity, privacy, security. *Information Polity 10*(1–2), 13–23. https://doi.org/10.3233/ip-2005-0066

Dutton, W. H., and Peltu, M. (1996). *Information and communication technologies: Visions and realities*. Oxford: Oxford University Press.

Dutton, W. H., and Shepherd, A. (2006). Trust in the internet as an experience technology. *Information, Communication & Society 9*(4), 433–451. https://doi.org/10.1080/13691180600858606

Dwoskin, E. (2018). Amazon is selling facial recognition to law enforcement—for a fistful of dollars. *Washington Post*, May 22.

Dwyer, C., Hiltz, S., and Passerini, K. (2007). Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. *AMCIS 2007 proceedings*, 339.

Earp, J. B., and Baumer, D. (2003). Innovative web use to learn about consumer behavior and online privacy. *Communications of the ACM 46*(4), 81–83. https://doi.org/10.1145/641205.641209

Eichenwald, K. (2013). Facebook leans in. *Vanity Fair*, May.

Ellison, N. B., Vitak, J., Steinfield, C., Gray, R., and Lampe, C. (2011). Negotiating privacy concerns and social capital needs in a social media environment. In S. Trepte and L. Reinecke (eds.), *Privacy online*, 19–32. Berlin: Springer.

Elueze, I., and Quan-Haase, A. (2018). Privacy attitudes and concerns in the digital lives of older adults: Westin's privacy attitude typology revisited. *American Behavioral Scientist 62*(10), 1372–1391. https://doi.org/10.1177/0002764218787026

Etzioni, A. (2007). Are new technologies the enemy of privacy? *Knowledge, Technology & Policy 20*(2), 115–119. https://doi.org/10.1007/s12130-007-9012-x

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.

Feagin, J., and Bennefield, Z. (2014). Systemic racism and US health care. *Social Science & Medicine 103*, 7–14. https://doi.org/10.1016/j.socscimed.2013.09.006

Fischer, C. S. (1994). *America calling: A social history of the telephone to 1940*. Berkeley: University of California Press.

Flaherty, D. H. (1989). *Protecting privacy in surveillance societies*. Chapel Hill: University of North Carolina Press.

Flaherty, K. (2018). This is why people no longer trust Google and Facebook with their data. Forbes. Oct 10, 2018.

Foucault, M. (1975). *Discipline and punish: The birth of the prison*. New York: Vintage.
Fuchs, C. (2015). *Culture and economy in the age of social media*. London: Routledge.
Gandy, O. H., Jr. (1989). The surveillance society: Information technology and bureaucratic social control. *Journal of Communication 39*(3), 61–76. https://doi.org/10.1111/j.1460-2466.1989.tb01040.x
Gandy, O. H., Jr. (1998). *Communication and race: A structural perspective*. London: Arnold.
Gandy, O. H., Jr. (2012). Statistical surveillance. In K. Ball, K. D. Haggerty, and D. Lyon (eds.), *Routledge handbook of surveillance studies*, 125–132. London: Routledge.
Gandy, O. H., Jr., and Nemorin, S. (2018). Toward a political economy of nudge: Smart city variations. *Information, Communication & Society 22*(14), 2112–2126. https://doi.org/10.1080/1369118X.2018.1477969
Gershgorn, D. (2018). Alexa is the future of Amazon's consumer business. *Quartz*, September 22.
Gibbs, J. L., Ellison, N. B., and Lai, C. H. (2011). First comes love, then comes Google: An investigation of uncertainty reduction strategies and self-disclosure in online dating. *Communication Research 38*(1), 70–100. https://doi.org/10.1177/0093650210377091
Giddens, A. (1983). Comments on the theory of structuration. *Journal for the Theory of Social Behaviour 13*(1), 75–80. https://doi.org/10.1111/j.1468-5914.1983.tb00463.x
Gillespie, T. (2014). Facebook's algorithm—why our assumptions are wrong, and our concerns are right. *Culture Digitally*, 4.
Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven, CT: Yale University Press.
Goffman, E. (1967). *Interaction ritual*. New York: Doubleday.
Hall, S. (2001). Encoding/decoding. In M. G. Durham and D. Kellner (eds.), *Media and cultural studies: Keyworks*, 137–144. Oxford: John Wiley & Sons.
Hall, W., and Pesenti, J. (2017). *Growing the artificial intelligence industry in the UK*. Independent Review for Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy. UK and the Commonwealth.
Hamilton, I. A. (2018). "I predict one day Amazon will fail. Amazon will go bankrupt": Jeff Bezos makes surprise admission about Amazon's life span. *Business Insider*, November 16.
Hamilton, J. T. (2000). *Channeling violence: The economic market for violent television programming*. Princeton, NJ: Princeton University Press.
Hamilton, J. T. (2004). *All the news that's fit to sell: How the market transforms information into news*. Princeton, NJ: Princeton University Press.
Hamilton, K., Karahalios, K., Sandvig, C., and Eslami, M. (2014). A path to understanding the effects of algorithm awareness. In *CHI'14 extended abstracts on human factors in computing systems*, 631–642. New York: ACM.
Hampton, K. N., and Hargittai, E. (2016). Stop blaming Facebook for Trump's election win. *The Hill*, November 23.
Hampton, K. N., and Wellman, B. (2018). Lost and saved . . . again: The moral panic

about the loss of community takes hold of social media. *Contemporary Sociology 47*(6), 643–651. https://doi.org/10.1177/0094306118805415

Hargittai, E. (2000). Radio's lessons for the internet. *Communications of the ACM 43*(1), 50–57. https://doi.org/10.1145/323830.323844

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday 7*(4). Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/942

Hargittai, E. (2006). Hurdles to information seeking: Spelling and typographical mistakes during users' online behavior. *Journal of the Association for Information Systems 7*(1), 52–67. https://doi.org/10.17705/1jais.00076

Hargittai, E. (2018). Potential biases in big data: Omitted voices on social media. *Social Science Computer Review 38*(1), 10–24. https://doi.org/10.1177/0894439318788322

Hargittai, E., and Hinnant, A. (2008). Digital inequality: Differences in young adults' use of the internet. *Communication Research 35*(5), 602–621. https://doi.org/10.1177/0093650208321782

Hargittai, E., and Marwick, A. (2016). "What can I really do?": Explaining the privacy paradox with online apathy. *International Journal of Communication 10*, 21. https://doi.org/1932-8036/20160005

Hasenclever, A., Mayer, P., and Rittberger, V. (1997). *Theories of international regimes*. Cambridge: Cambridge University Press.

Heidegger, M. (1962). *Being and time*. New York: Harper.

Helsper, E. J. (2017). The social relativity of digital exclusion: Applying relative deprivation theory to digital inequalities. *Communication Theory 27*(3), 223–242. https://doi.org/10.1111/comt.12110

Hintz, A., Dencik, L., and Wahl-Jorgensen, K. (2017). Digital citizenship and surveillance | digital citizenship and surveillance society—introduction. *International Journal of Communication 11*, 9. https://doi.org/1932-8036/20170005

Ho, A., Hancock, J., and Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication 68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Horwitz, R. B. (1991). *The irony of regulatory reform: The deregulation of American telecommunications*. Oxford: Oxford University Press.

Howard, P. N., and Jones, S. (2004). *Society online: The internet in context*. New York: Sage.

Huesmann, L. R., and Taylor, L. D. (2003). The case against the case against media violence. In D. A. Gentile (ed.), *Advances in applied developmental psychology: Media violence and children: A complete guide for parents and professionals*, 107–130. Westport, CT: Praeger Publishers / Greenwood Publishing Group.

Hughes, T. P. (1987). The evolution of large technological systems. In W. Bijker, T. P. Hughes, and T. Pinch (eds.), *The social construction of technological systems: New directions in the sociology and history of technology*, 51–82. Cambridge, MA: MIT Press.

Humphreys, L. (2011). Who's watching whom? A study of interactive technology and surveillance. *Journal of Communication 61*(4), 575–595. https://doi.org/10.1111/j.1460-2466.2011.01570.x

Humphreys, L. (2018). *The qualified self: Social media and the accounting of everyday life.* Cambridge, MA: MIT Press.

Humphreys, L., Gill, P., and Krishnamurthy, B. (2010). How much is too much? Privacy issues on Twitter. Conference of International Communication Association, Singapore, June. Retrieved from https://people.cs.umass.edu/~phillipa/papers/ica10.pdf

Hypponen, M. (2013). How the NSA betrayed the world's trust-time to act. TED. November 7.

Inglehart, R., and Wezel, C. (2005). *Modernization, cultural change, and democracy: The human development sequence.* Cambridge: Cambridge University Press.

Ittelson, W. H., et al. (1970). The use of behavioural maps in environmental psychology. In H. M. Prohansky, W. H. Ittelson, and L. G. Rivlin (eds.), *Environmental psychology: Man and his physical setting*, 658–668. New York: Rinehart & Winston.

Jackson, S. J. (2014). Rethinking repair. In T. Gillespie, P. Boczkowski, and K. Foot (eds.), *Media technologies: Essays on communication, materiality, and society*, 221–239. Cambridge, MA: MIT Press.

Jai, T. M. C., and King, N. J. (2016). Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers? *Journal of Retailing and Consumer Services 28*, 296–303. https://doi.org/10.1016/j.jretconser.2015.01.005

Jiang, L. C., Bazarova, N. N., and Hancock, J. T. (2011). The disclosure-intimacy link in computer-mediated communication: An attributional extension of the hyperpersonal model. *Human Communication Research 37*(1), 58–77. https://doi.org/10.1111/j.1468-2958.2010.01393.x

Jones, S. (1998). *Doing internet research: Critical issues and methods for examining the Net*. New York: Sage Publications.

Kahneman, D., and Egan, P. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kang, J. (1998). Information privacy in cyberspace transactions. *Stanford Law Review 50*, 1193–1294. Retrieved from https://www.ntia.doc.gov/legacy/ntiahome/privacy/files/CPRIVACY.PDF

Katz, J. E., and Rice, R. E. (2009). Public views of mobile medical devices and services: A US national survey of consumer sentiments towards RFID healthcare technology. *International Journal of Medical Informatics 78*(2), 104–114. https://doi.org/10.1016/j.ijmedinf.2008.06.001

Kozlowska, H., Gershgorn, D., and Todd, S. (2018). The Cambridge Analytica scandal is wildly confusing. This timeline will help. *Quartz*, March 29. Retrieved from https://qz.com/1240039/the-cambridge-analytica-scandal-is-confusing-this-timeline-will-help/

Kramer, A. D., Guillory, J. E., and Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences 111*(24), 8788–8790. https://doi.org/10.1073/pnas.1320040111

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Princeton, NJ: Princeton University Press.

Kwak, N. (1999). Revisiting the knowledge gap hypothesis: Education, motivation, and media use. *Communication Research 26*(4), 385–413. https://doi.org/10.1177/009365099026004002

Langheinrich, M. (2001). Privacy by design—principles of privacy-aware ubiquitous systems. In *International conference on ubiquitous computing*, 273–291. Berlin: Springer.

LaRose, R., and Rifon, N. (2006). Your privacy is assured—of being disturbed: Websites with and without privacy seals. *New Media & Society 8*(6), 1009–1029. https://doi.org/10.1177/1461444806069652

Lessig, L. (2009). *Code: And other laws of cyberspace*. New York: Vintage.

Li, T., and Unger, T. (2012). Willing to pay for quality personalization? Trade-off between quality and privacy. *European Journal of Information Systems 21*(6), 621–642. https://doi.org/10.1057/ejis.2012.13

Ling, R., Bjelland, J., Sundsøy, P. R., and Campbell, S. W. (2014). Small circles: Mobile telephony and the cultivation of the private sphere. *Information Society 30*(4), 282–291. https://doi.org/10.1080/01972243.2014.915279

Ling, R., and Lai, C. H. (2016). Microcoordination 2.0: Social coordination in the age of smartphones and messaging apps. *Journal of Communication 66*(5), 834–856. https://doi.org/10.1111/jcom.12251

Litt, E. (2013). Measuring users' internet skills: A review of past assessments and a look toward the future. *New Media & Society 15*(4), 612–630. https://doi.org/10.1177/1461444813475424

Litt, E., and Hargittai, E. (2014). Smile, snap, and share? A nuanced approach to privacy and online photo-sharing. *Poetics 42*, 1–21. https://doi.org/10.1016/j.poetic.2013.10.002

Livingstone, S., Mascheroni, G., and Staksrud, E. (2018). European research on children's internet use: Assessing the past and anticipating the future. *New Media & Society 20*(3), 1103–1122. https://doi.org/10.1177/1461444816685930

Lloyd, M. (2010). *Prologue to a farce: Communication and democracy in America*. Urbana: University of Illinois Press.

Lupton, D. (2013). Quantifying the body: Monitoring and measuring health in the age of mHealth technologies. *Critical Public Health 23*(4), 393–403. https://doi.org/10.1080/09581596.2013.794931

Lupton, D. (2016). *The quantified self*. London: John Wiley & Sons.

Lutz, C., Hoffman, C. P., Bucher, E., and Fieseler, C. (2018). The role of privacy concerns in the sharing economy. *Information, Communication & Society 21* (10), 1472–1492. https://doi.org/10.1080/1369118X.2017.1339726

Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society 1*(2). https://doi.org/10.1177/2053951714541861

Madden, M. (2014). Public Perceptions of Privacy and Security in the Post-Snowden Era. Pew Internet, November 12. Retrieved from https://www.pewresearch.org/internet/2014/11/12/public-privacy-perceptions/

Madden, M., and Raine, L. (2015). Americans' Attitudes about Privacy, Security and Surveillance. Pew Internet, May 20. Retrieved from https://www.pewresearch.org/internet/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

Madden, M. and Smith, A. (2010). Reputation Management and Social Media. Pew Internet, May 26. Retrieved from https://www.pewresearch.org/internet/2010/05/26/reputation-management-and-social-media/

Marwick, A. E., and boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society 16*(7), 1051–1067. https://doi.org/10.1177/1461444814543995

Marwick, A., and Hargittai, E. (2018). Nothing to hide, nothing to lose? Incentives and disincentives to sharing information with institutions online. *Information, Communication & Society*, 1–17. https://doi.org/10.1080/1369118X.2018.1450432

Marx, G. T. (1998). Ethics for the new surveillance. *Information Society 14*(3), 171–185. https://doi.org/10.1080/019722498128809

Marx, K. (1867). *Capital*. Vol. 1. Mineola, NY: Courier Dover Publications.

Masur, P. K. (2018). *Situational privacy and self-disclosure: Communication processes in online environments*. Berlin: Springer.

May, C., and Finch, T. (2009). Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology 43*(3), 535–554. https://doi.org/10.1177/0038038509103208

McKinnon, J., and MacMillan, D. (2018). Google says it continues to allow apps to scan data from Gmail accounts. *Wall Street Journal*, September 20.

McLuhan, M., and Fiore, Q. (1967). *The medium is the message*. Berkeley, CA: Gingko Press.

Metzger, M. J. (2007). Communication privacy management in electronic commerce. *Journal of Computer-Mediated Communication 12*(2), 335–361. https://doi.org/10.1111/j.1083-6101.2007.00328.x

Meyer, R. (2014). Everything we know about Facebook's secret mood manipulation experiment. *The Atlantic*, June 28.

Michels, R. (1915). *Political parties: A sociological study of the oligarchical tendencies of modern democracy*. New York: Hearst's International Library Company.

Miller, A. R. (1971). *The assault on privacy: Computers, data banks, and dossiers*. Ann Arbor: University of Michigan Press.

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research 26*(4), 323–339. https://doi.org/10.1086/209566

Mosco, V. (2015). *To the cloud: Big data in a turbulent world*. London: Routledge.

Mueller, M. L. (2009). *Ruling the root: Internet governance and the taming of cyberspace*. Cambridge, MA: MIT Press.

Napoli, P. M. (1999). Deconstructing the diversity principle. *Journal of Communication 49*(4), 7–34. https://doi.org/10.1111/j.1460-2466.1999.tb02815.xCitations

Napoli, P. M. (2001). *Foundations of communications policy: Principles and process in the regulation of electronic media*. New York: Hampton Press.

Napoli, P. M. (2011). Exposure diversity reconsidered. *Journal of Information Policy 1*, 246–259. https://doi.org/10.5325/jinfopoli.1.2011.0246

Napoli, P. M. (2012). *Audience economics: Media institutions and the audience marketplace.* New York: Columbia University Press.

Napoli, P. M. (2015). Social media and the public interest: Governance of news plat-

forms in the realm of individual and algorithmic gatekeepers. *Telecommunications Policy 39*(9), 751–760. https://doi.org/10.1016/j.telpol.2014.12.003

Napoli, P. M., and Caplan, R. (2017). Why media companies insist they're not media companies, why they're wrong, and why it matters. *First Monday 22*(5). Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/7051/6124

Neuman, W. R. (1986). *The paradox of mass politics*. Cambridge, MA: Harvard University Press.

Neuman, W. R. (1991). *The future of the mass audience*. London: Cambridge University Press.

Neuman, W. R. (2010). *Media, technology, and society: Theories of media evolution*. Ann Arbor: University of Michigan Press.

Neuman, W. R. (2016). *The digital difference*. Cambridge, MA: Harvard University Press.

Neuman, W. R., Bimber, B., and Hindman, M. (2011). The internet and four dimensions of citizenship. In G. C. Edwards, L. R. Jacobs, and Robert Y. Shapiro (eds.), *The Oxford handbook of American public opinion and the media*, 22–42. Oxford: Oxford University Press.

Neuman, W. R., Guggenheim, L., Mo Jang, S., and Bae, S. Y. (2014). The dynamics of public attention: Agenda-setting theory meets big data. *Journal of Communication 64*(2), 193–214. https://doi.org/10.1111/jcom.12088

Neuman, W. R., McKnight, L., and Solomon, R. J. (1997). *The Gordian knot: Political gridlock on the information highway*. Cambridge, MA: MIT Press.

Neuman, W. R., Park, Y. J., and Panek, E. (2012). Info capacity| Tracking the flow of information into the home: An empirical assessment of the digital revolution in the US from 1960–2005. *International Journal of Communication 6*, 20. https://doi.org/1932-8036/20121022

New America Foundation (2017). Ranking Digital Rights Partners with Consumer Reports to Set Standards for Privacy and Security: New Digital Standard will Safeguard Consumer Privacy in the Digital Marketplace. Press Release, March 6. Retrieved from https://thedigitalstandard.org

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review 79*, 119. Retrieved from https://crypto.stanford.edu/portia/papers/RevnissenbaumDTP31.pdf

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Palo Alto, CA: Stanford University Press.

Nix, N., House, B., and Allison, B. (2018). Facebook goes on a hiring spree for Washington lobbyists. *Boomerang*, March 27.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: NYU Press.

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs 41*(1), 100–126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

Olivero, N., and Lunt, P. (2004). Privacy versus willingness to disclose in e-commerce exchanges: The effect of risk awareness on the relative role of trust and control. *Journal of Economic Psychology 25*(2), 243–262. https://doi.org/10.1016/S0167-4870(02)00172-1

Orwell, G. (1949). *Nineteen eighty-four*. London: Gardners Books.
Owen, B. M. (1978). The economic view of programming. *Journal of Communication 28*(2), 43–47. https://doi.org/10.1111/j.1460-2466.1978.tb01592.x
Park, Y. J. (2008). Privacy regime, culture and user practices in the cyber-marketplace. *Info 10*(2), 57–74. https://doi.org/10.1108/14636690810862811
Park, Y. J. (2011a). Market philosophy and information privacy. *Javnost—The Public 18*(2), 87–99. https://doi.org/10.1080/13183222.2011.11009058
Park, Y. J. (2011b). Provision of internet privacy and market conditions: An empirical analysis. *Telecommunications Policy 35*(7), 650–662. https://doi.org/10.1016/j.telpol.2011.06.003
Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research 40*(2), 215–236. https://doi.org/10.1177/0093650211418338
Park, Y. J. (2018a). Explicating net diversity in trend assessment. *Communication Research 45*(5), 783–809. https://doi.org/10.1177/0093650215601883
Park, Y. J. (2018b). Social antecedents and consequences of political privacy. *New Media & Society 20*(7), 2352–2369. https://doi.org/10.1177/1461444817716677
Park, Y. J. (under review). Government surveillance and citizen reactance.
Park, Y. J., and Chung, J. E. (2017). Health privacy as sociotechnical capital. *Computers in Human Behavior 76*, 227–236. https://doi.org/10.1016/j.chb.2017.07.025
Park, Y. J., Chung, J. E., and Shin, D. H. (2018). The structuration of digital ecosystem, privacy, and big data intelligence. *American Behavioral Scientist 62*(10), 1319–1337. https://doi.org/10.1177/0002764218787863
Park, Y. J., and Jang, S. M. (2014). Understanding privacy knowledge and skill in mobile communication. *Computers in Human Behavior 38*, 296–303. https://doi.org/10.1016/j.chb.2014.05.041
Park, Y. J., and Jang, S. M. (2017). Public attention, social media, and the Edward Snowden saga. *First Monday 22*(8). Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/7818
Park, Y. J., Jang, S. M., Lee, H., and Yang, G. S. (2018). Divide in Ferguson: Social media, social context, and division. *Social Media + Society 4*(3), https://doi.org/10.1177/2056305118789630
Park, Y. J., and Shin, D. (2020). Contextualizing privacy on health-related use of information technology. *Computers in Human Behavior 105* (April), 106204. https://doi.org/10.1016/j.chb.2019.106204
Park, Y. J., and Skoric, M. (2017). Personalized ad in your Google Glass? Wearable technology, hands-off data collection, and new policy imperative. *Journal of Business Ethics 142*(1), 71–82. https://doi.org/10.1007/s10551-015-2766-2
Park, Y. J., and Yang, G. S. (2017). Personal network on the internet: How the socially marginalized stay marginalized in personal network diversity and multiplicity. *Telematics and Informatics 34*(1), 1–10. https://doi.org/10.1016/j.tele.2016.04.001
Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.
Patil, S., Norcie, G., Kapadia, A., and Lee, A. J. (2012). Reasons, rewards, regrets: Privacy considerations in location sharing as an interactive practice. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, 5. New York: ACM.

Pearce, K. E., and Rice, R. E. (2013). Digital divides from access to activities: Comparing mobile and personal computer internet users. *Journal of Communication 63*(4), 721–744. https://doi.org/10.1111/jcom.12045

Pearson, S., and Charlesworth, A. (2009). Accountability as a way forward for privacy protection in the cloud. In *IEEE International Conference on Cloud Computing* (131–144). Berlin: Springer.

Pedersen, D. M. (1999). Model for types of privacy by privacy functions. *Journal of Environmental Psychology 19*(4), 397–405. https://doi.org/10.1006/jevp.1999.0140

Petronio, S. (1991). Communication boundary management: A theoretical model of managing disclosure of private information between marital couples. *Communication Theory 1*(4), 311–335. https://doi.org/10.1111/j.1468-2885.1991.tb00023.x

Petronio, S. (2010). Communication privacy management theory: What do we know about family privacy regulation? *Journal of Family Theory & Review 2*(3), 175–196. https://doi.org/10.1111/j.1468-2885.1991.tb00023.x.

Petronio, S. (2012). *Boundaries of privacy: Dialectics of disclosure*. Albany: SUNY Press.

Phelps, J. E., D'Souza, G., and Nowak, G. J. (2001). Antecedents and consequences of consumer privacy concerns: An empirical investigation. *Journal of Interactive Marketing 15*(4), 2–17. https://doi.org/10.1504/IJEMR.2011.045610

Pickard, V. (2013). Social democracy or corporate libertarianism? Conflicting media policy narratives in the wake of market failure. *Communication Theory 23*(4), 336–355. https://doi.org/10.1111/comt.12021

Pickard, V. (2018). The strange life and death of the fairness doctrine: Tracing the decline of positive freedoms in American policy discourse. *International Journal of Communication 12,* 3434–3453. https://doi.org/1932-8036/20180005

Pinch, T. J. (2010). The invisible technologies of Goffman's sociology from the merry-go-round to the internet. *Technology and Culture 51*(2), 409–424. https://doi.org/10.1353/tech.0.0456

Pinch, T. J., and Bijker, W. E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science 14*(3), 399–441. https://doi.org/10.1177/030631284014003004

Poltash, N. A. (2012). Snapchat and sexting: A snapshot of baring your bare essentials. *Richmond Journal of Law and Technology 19*, 1. Retrieved from https://scholarship.richmond.edu/jolt/vol19/iss4/3/

Pool, I. d. S. (1983a). *Technologies of freedom*. Cambridge, MA: Harvard University Press.

Pool, I. d. S. (1983b). *Forecasting the telephone: A retrospective technology assessment.* New York: Ablex Publication.

Popescu, M., Baruh, L., Messaris, P., and Humphreys, L. (2017). Consumer surveillance and distributive privacy harms in the age of big data. In *Digital media: Transformations in human communication*, 313–327. New York: Peter Lang.

Posner, R. A. (1981). The economics of privacy. *American Economic Review 71*(2), 405–409. Retrieved from https://www.jstor.org/stable/1815754?seq=1#metadata_info_tab_contents

Post, R. C. (2000). Three concepts of privacy. *Georgetown Law Journal 89*, 2087–

2098. Retrieved from https://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?article=1184&context=fss_papers

Rains, S. A., and Bosch, L. A. (2009). Privacy and health in the information age: A content analysis of health web site privacy policy statements. *Health Communication 24*(5), 435–446. https://doi.org/10.1080/10410230903023485

Rathi, R. (2018). Deep learning and "hyper-personalization" are the future of marketing automation. *Entrepreneur*, August 15.

Reidenberg, J. R. (2001). E-commerce and trans-Atlantic privacy. *Houston Law Review 38*, 717–749. Retrieved from https://houstonlawreview.org/article/4080-e-commerce-and-trans-atlantic-privacy

Reisdorf, B., Hampton, K., Fernandez, L., and Dutton, W. H. (2018). Broadband to the neighborhood: Digital divides in Detroit. Retrieved from https://www.researchgate.net/publication/324665092_Broadband_to_the_Neighborhood_Digital_Divides_in_Detroit

Resnick, R. (2002). Beyond bowling together: Socio technical capital. In J. Carroll (ed.), *Human-computer interaction in the new millennium*, 242–272. Reading, MA: Addison-Wesley.

Ribak, R., and Turow, J. (2003). Internet power and social context: A globalization approach to web privacy concerns. *Journal of Broadcasting & Electronic Media 47*(3), 328–349. https://doi.org/10.1207/s15506878jobem4703_2

Rice, R. E., and Katz, J. E. (2003). Comparing internet and mobile phone usage: Digital divides of usage, adoption, and dropouts. *Telecommunications Policy 27*(8–9), 597–623. https://doi.org/10.1016/S0308-5961(03)00068-5

Rotenberg, M. (2001). Fair information practices and the architecture of privacy (what Larry doesn't get). *Stanford Technology Law Review 1*. Retrieved from https://www.bibsonomy.org/bibtex/571c2cdb6b3edd84b3a44bd1ab1d1920

Sadowski, J., and Pasquale, F. A. (2015). The spectrum of control: A social theory of the smart city. *First Monday 20*(7). Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/5903/4660

Samavi, R., and Consens, M. P. (2018). Publishing privacy logs to facilitate transparency and accountability. *Journal of Web Semantics 50*, 1–20. https://doi.org/10.1016/j.websem.2018.02.001

Sandvig, C., Hamilton, K., Karahalios, K., and Langbort, C. (2015). Can an algorithm be unethical? Ann Arbor, 1001(48109), 1285. Retrieved from http://social.cs.uiuc.edu/papers/pdfs/ICA2015-Sandvig.pdf

Sandvig, C., Hamilton, K., Karahalios, K., and Langbort, C. (2016). Automation, algorithms, and politics | when the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication 10*, 19. Retrieved from 1932-8036/20160005

Sapolsky, R. M. (2004). Social status and health in humans and other animals. *Annual Review Anthropology 33*, 393–418. https://doi.org/10.1146/annurev.anthro.33.070203.144000

Schoeman, F. (1984). Privacy: Philosophical dimensions. *American Philosophical Quarterly 21*(3), 199–213. https://doi.org/http://hdl.handle.net/10822/804170

Schudson, M. (2013). *Advertising, the uneasy persuasion: Its dubious impact on American society.* New York: Routledge.

Schwartz, B. (1968). The social psychology of privacy. *American Journal of Sociology 73*(6), 741–752. https://doi.org/10.1086/224567

Schwartz, J. (2001). Giving web a memory costs its user privacy. *New York Times*, September 4.

Scola, N. (2018). Google is building a city of the future in Toronto. Would anyone want to live there? *Politico*, July–August.

Sellar, S., and Thompson, G. (2016). The becoming-statistic: Information ontologies and computerized adaptive testing in education. *Cultural Studies–Critical Methodologies 16*(5), 491–501. https://doi.org/10.1177/1532708616655770

Shapiro, C., and Varian, H. R. (1998). *Information rules: A strategic guide to the network economy.* Cambridge, MA: Harvard Business Press.

Shannon, C. E., and Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Singer, E., Mathiowetz, N. A., and Couper, M. P. (1993). The impact of privacy and confidentiality concerns on survey participation the case of the 1990 US census. *Public Opinion Quarterly 57*(4), 465–482. https://doi.org/10.1177/1532673X9802600104

Solove, D. J. (2001). Privacy and power: Computer databases and metaphors for information privacy. *Stanford Law Review 53*, 1393–1462. Retrieved from https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2077&context=faculty_publications

Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist 43*(3), 377–391. https://doi.org/10.1177/00027649921955326

Steiner, P. O. (1952). Program patterns and preferences, and the workability of competition in radio broadcasting. *Quarterly Journal of Economics 66*(2), 194–223. https://doi.org/10.2307/1882942

Stiglitz, J. E. (2003). Globalization and growth in emerging markets and the new economy. *Journal of Policy Modeling 25*(5), 505–524. https://doi.org/10.1016/S0161893803)00043-7

Streeter, T. (1996). *Selling the air: A critique of the policy of commercial broadcasting in the United States*. Chicago: University of Chicago Press.

Sunstein, C. R. (2018). *# Republic: Divided democracy in the age of social media*. Princeton, NJ: Princeton University Press.

Tidwell, L. C., and Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research 28*(3), 317–348. https://doi.org/10.1111/j.1468-2958.2002.tb00811.x

Tönnies, F. (1887). Community and society. In J. Lin and C. Mele (eds.), *The urban sociology reader*, 13. New York: Psychology Press & Routledge Classic Editions.

TRUSTe. (2008). Consumer attitudes about behavioral targeting. Research conducted by TNS Global, March. Retrieved from http://danskprivacynet.files.wordpress.com/2009/02/truste2008_tns_bt_study_summary1.pdf

Turow, J. (2003). *Americans and online privacy: The system is broken*. A Report from the Annenberg Public Policy Center of the University of Pennsylvania. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1411&context=asc_papers

Turow, J. (2005). Audience construction and culture production: Marketing surveillance in the digital age. *Annals of the American Academy of Political and Social Science 597*(1), 103–121. https://doi.org/10.1177/0002716204270469

Turow, J. (2017). *The aisles have eyes: How retailers track your shopping, strip your privacy, and define your power*. New Haven, CT: Yale University Press.

Turow, J., Delli Carpini, M. X., Draper, N. A., and Howard-Williams, R. (2012). Americans roundly reject tailored political advertising—at a time when political campaigns are embracing it. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1414&context=asc_papers

Turow, J., Hennessy, M., and Draper, N. (2015). The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1554&context=asc_papers

Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science 211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Tversky, A., and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business 59*(4), S251–S278. https://doi.org/0021-939818615904-001

Utz, S., and Krämer, N. C. (2009). The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Journal of Psychosocial Research on Cyberspace 3*(2). Retrieved from https://cyberpsychology.eu/article/view/4223/3265

Van Deursen, A., and Van Dijk, J. (2011). Internet skills and the digital divide. *New Media & Society 13*(6), 893–911. https://doi.org/10.1177/1461444810386774

Van Dijk, J. (2012). *The network society*. New York: Sage Publications.

Venturelli, S. (2002). Inventing e-regulation in the US, EU and East Asia: Conflicting social visions of the information society. *Telematics and Informatics 19*(2), 69–90. https://doi.org/10.1016/S0736-5853(01)00007-7

Vogel, H. L. (2014). *Entertainment industry economics: A guide for financial analysis*. New York: Cambridge University Press.

Vogel, S. K. (1998). *Freer markets, more rules: Regulatory reform in advanced industrial countries.* Ithaca, NY: Cornell University Press.

Warren, S. D., and Brandeis, L. D. 1890. The right to privacy. *Harvard Law Review 4*(5), 193–220.

Webster, J. G., and Ksiazek, T. B. (2012). The dynamics of audience fragmentation: Public attention in an age of digital media. *Journal of Communication 62*(1), 39–56. https://doi.org/10.1111/j.1460-2466.2011.01616.x

Westin, A. F. (1984). *The origins of modern claims to privacy*. Cambridge: Cambridge University Press.

Westin, A. F. (1998). *Ecommerce and privacy: What net users want*. Technical Report for Privacy and American Business and PricewaterhouseCoopers. Hackensack, NJ: Privacy & American Business.

Westin, A. F. (2001). *Privacy on & off the internet: What consumers want*. Technical

Report for Privacy & American Business. Hackensack, NJ: Privacy & American Business.

Westin, A. F. (2003). Social and political dimensions of privacy. *Journal of Social Issues 59*(2), 431–453. https://doi.org/10.1111/1540-4560.00072

Wong, N. (2019). Backlash against tech giants hindering Sidewalk Labs' Toronto plan, executive says: Sidewalk executive says company doesn't work closely with Google in any ongoing way other than on some very specific projects. Financial Post, *Bloomberg News*, July 2.

Wu, T. (2017). *The attention merchants: The epic scramble to get inside our heads*. New York: Vintage.

Xu, H., Dinev, T., Smith, J., and Hart, P. (2011). Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems 12*(12), 798. https://doi.org/10.17705/1jais.00281

Yan, M. Z., and Park, Y. J. (2009). Duopoly ownership and local informational programming on broadcast television: Before–after comparisons. *Journal of Broadcasting & Electronic Media 53*(3), 383–399. https://doi.org/10.1080/08838150903102709

Yanich, D. (2013). Local TV news, content, and the bottom line. *Journal of Urban Affairs 35*(3), 327–342. https://doi.org/10.1111/j.1467-9906.2012.00637.x

Yao, M. Z., Rice, R. E., and Wallis, K. (2007). Predicting user concerns about online privacy. *Journal of the American Society for Information Science and Technology 58*(5), 710–722. https://doi.org/10.1002/asi.20530

Zillien, N., and Hargittai, E. (2009). Digital distinction: Status-specific types of internet usage. *Social Science Quarterly 90*(2), 274–291. https://doi.org/10.1111/j.1540-6237.2009.00617.x

Zimmer, M. (2008). The externalities of search 2.0: The emerging privacy threats when the drive for the perfect search engine meets Web 2.0. *First Monday 13*(3). Retrieved from https://journals.uic.edu/ojs/index.php/fm/article/view/2136/1944

Zuboff, S. (2019). *The age of surveillance capitalism*. London: Profile.

Index