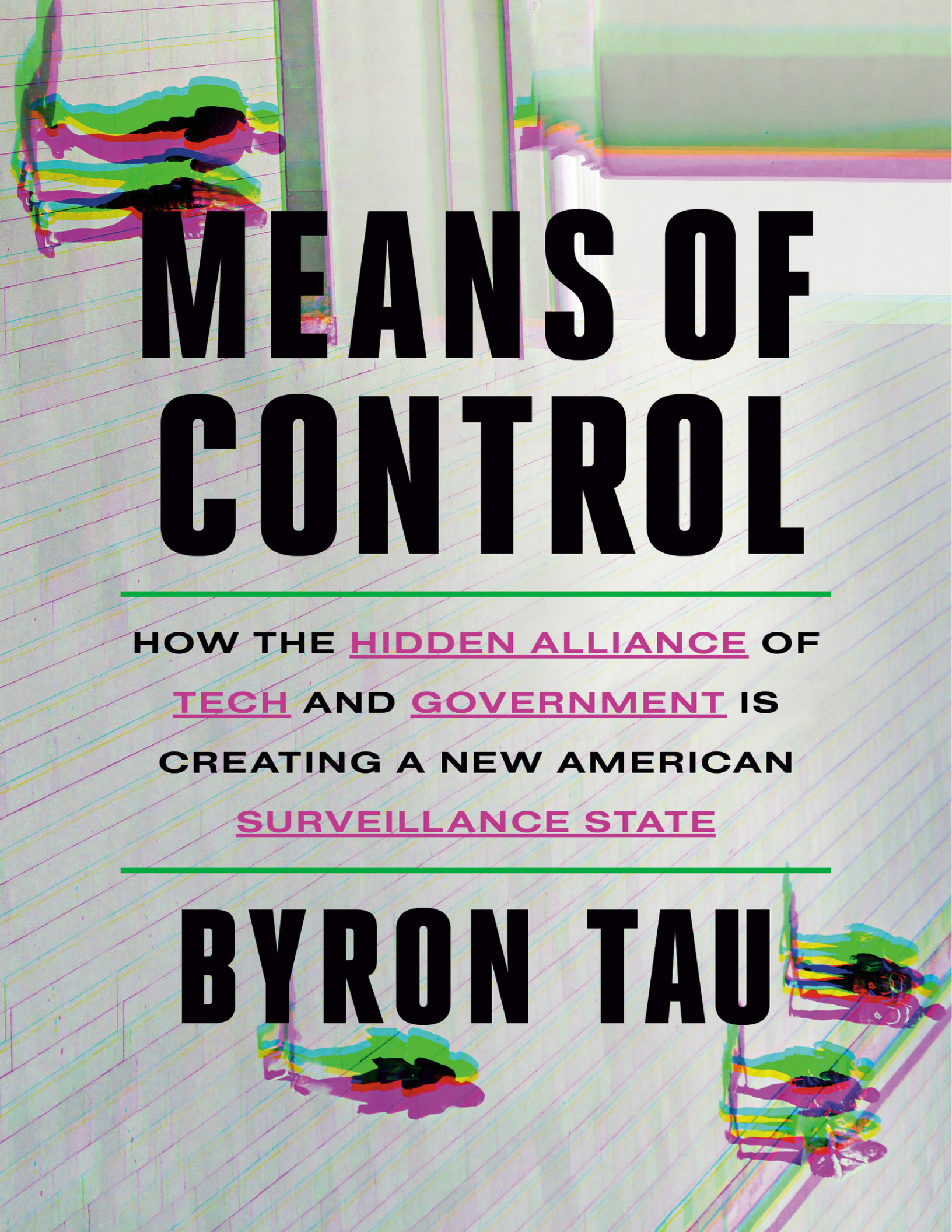

MEANS OF CONTROL

HOW THE HIDDEN ALLIANCE OF TECH AND GOVERNMENT IS CREATING A NEW AMERICAN SURVEILLANCE STATE

BYRON TAU

MEANS OF CONTROL

HOW THE HIDDEN ALLIANCE OF TECH AND GOVERNMENT IS CREATING A NEW AMERICAN SURVEILLANCE STATE

BYRON TAU

CROWN
NEW YORK

Published in the United States by Crown, an imprint of the Crown Publishing Group, a division of Penguin Random House LLC, New York.

Hardback ISBN 9780593443224
Ebook ISBN 9780593443231

crownpublishing.com

Book design by Elizabeth A. D. Eno, adapted for ebook

Cover design: Yang Kim
Cover photograph: Busà Photography / Getty Images

ep_prh_6.3_146236082_c0_r0

Once the technical means of control have reached a certain size, a certain degree of being connected one to another, the chances for freedom are over for good. The word has ceased to have meaning.

—Thomas Pynchon, Gravity's Rainbow

And in this sense, all Americans are Marxists, for we believe nothing if not that history is moving us toward some preordained paradise and that technology is the force behind that movement.

—Neil Postman, Amusing Ourselves to Death

CONTENTS

AUTHOR'S NOTE

What follows is a work of nonfiction. Like all journalism, it's the best possible approximation of the truth. It is based on more than 350 interviews and tens of thousands of pages of documents. The narrative is supported with extensive documentation in the endnotes whenever possible. But because this is a book about intelligence and law enforcement—in some cases, classified intelligence or sensitive law enforcement methods—a full recounting of every source is not possible. Most people spoke to me only on the condition that their words not be attributed to them. The U.S. government has not always taken kindly to its employees and contractors speaking to the media and has often found ways to visit professional or reputational consequences on those who do. When sources have asked to remain anonymous, I have always tried to weigh the motivations of these sources; to evaluate what they have told me with a skeptical eye; and to seek corroboration for anything and everything in any way I can.

For those willing to speak on the record, I am eternally grateful. But I'm also grateful for those willing to speak at all. Journalists and intelligence officers are alike in some ways—except spies ask their foreign recruits to commit treason and reporters ask their sources to commit acts of transparency. Most people who cooperated seemed to do so out of a belief that the public should know more—more about their government's activities, more about their technology, and more about a shadowy industry that arose with scant public notice or debate. At the same time, my highest obligation is to the truth as I found it. Not everyone who appears in these pages will like what's said about them. I pulled no punches and did no favors.

The reader should not assume that because a person appears in this narrative, they were a source for this book. In many cases, people in the following pages refused to cooperate, and what is attributed to them was drawn from written accounts, documents, or other people familiar with the events described. Even when someone has refused to speak to me, I hav sought to be fair, to seek their perspective in any way I can, to try to se things their way, and not to assume the worst about them or their actions.

I have tried whenever possible to cite emails or documents. I have tried to take copious and contemporaneous notes. A small amount of dialogue is reconstruction from memory. I have sometimes lightly cleaned up quotations to fix grammar or syntax but never in a way that altered the meaning of the quotation. I have tried to visit locations in person whenever possible. I have very occasionally based short anecdotes or fleeting scenes on a single account, but only when it comported broadly with what could be checked and only when the source had an extensive track record of being reliable. Nothing of significance in this book is based on a single source. Every major claim i supported by numerous sources and documents, oftentimes dozens of sources and thousands of pages.

I have aimed to be transparent with the reader about what I do not know, and I have always tried to approach every story with an open mind and with the possibility that I am wrong. As the reporter and press critic Jack Shafer once said, a journalist should "follow a hunch with reporting that could undermine the hunch, address possible criticisms, remain open to criticism and refutation, correct meaningful errors of fact, abandon dry wells instead of pretending they're gushers."

One note on the use of the word "anonymized" as it relates to data sets. Documents or people may be quoted or paraphrased saying that bulk dat sets were "anonymized," or stripped of personal information. As a factual matter, I dispute in most cases that data at issue can truly be "anonymized" and have rarely, if ever, used that characterization in my own writing. Stripping data of personal information and replacing it with a random identifier should be properly described as "pseudonymization." There is

ample evidence that individuals can be reidentified in nearly all of the data sets covered in this book.

One of the challenges of reporting a complex story filled with technical details about technology and government is keeping track of acronyms and language that comes across as jargon. I've done my best to explain the technical developments covered in the book in plain English. I've also included a list of key concepts and definitions in the back of the book, and I encourage you to consult it frequently.

As a matter of full disclosure, my partner worked for a short time as a lawyer representing a company, Booz Allen Hamilton, that is briefly mentioned in this text. Her work was on an antitrust matter that is not at issue in this text or any of my other reporting. She has never provided me with nonpublic information regarding her work, nor has she been a source for this book or any of my other journalism. A final note to disclose: Until the fall of 2023, I was employed by The Wall Street Journal, whose parent company Dow Jones, owns Factiva, which is a competitor to some of the data brokers mentioned in a very limited sense. However, Factiva largely competes against those brokers in the media monitoring and corporate research verticals and does not have stores of advertising or consumer data, the subject of this book.

Finally, journalism is a human endeavor, and as in all human endeavors perfection is elusive. All mistakes are my responsibility.

INTRODUCTION

The Grindr Problem and a Wine-Soaked Dinner

In 2019, a government contractor and technologist named Mike Yeagley began making the rounds in Washington with a blunt warning for anyone in the country's national security establishment who would listen: the U.S. government had a Grindr problem.

A popular dating and hookup app, Grindr had launched ten years prior and had become a sensation. It relied on the GPS capabilities of modern smartphones to connect potential partners through the app—bringing together users in the same city, neighborhood, or even building. The app can show how far away a potential partner is in real time, down to the foot.

The app quickly amassed millions of users and became an essential part of gay culture around the globe. As Tom Capon, a young gay man who came of age just as the app was coming on the scene, put it in a 2019 essay, "It's no longer necessary to head to a gay bar to try your luck."

But to Yeagley, the app was something else: one of the tens of thousands of carelessly designed mobile phone apps that leaked massive amounts of data into the opaque world of online advertisers. That data, Yeagley knew, was easily accessible by anyone with a little technical know-how. So Yeagley—a

technology consultant in his late forties who had worked in and around government projects nearly his entire career—made a PowerPoint presentation and went on a road show around Washington to demonstrate precisely how that data was a serious national security risk.

As he would explain in a succession of bland government conference rooms, Yeagley was able to access the geolocation data on Grindr users through a hidden but ubiquitous entry point: the digital advertising exchanges that serve up the little digital banner ads along the top of not only Grindr but nearly every ad-supported mobile app and website. This was possible because a good chunk of the online ad space in the world was sold through near instantaneous auctions in a process called real-time bidding and those auctions were rife with surveillance potential. You know that ad that seems to be following you around the internet? Well, it's tracking you in more ways than one. In some cases, it's making your precise location available in near–real time to both advertisers and people like Mike Yeagley, who specialized in obtaining unique data sets for government agencies.

Working with Grindr data, Yeagley began drawing what are called geofences around buildings belonging to government agencies that do national security work. He was looking for phones belonging to Grindr users who spent their daytime hours at government office buildings. If the device spent most workdays at the Pentagon, the FBI headquarters, or the National Geospatial-Intelligence Agency (NGA) building at Fort Belvoir, for example, there was a good chance its owner worked for one of those agencies. Then he started looking at the movement of those phones through the Grindr data. When they weren't at their offices, where did they go? A small number of them had lingered at highway rest stops in the D.C. area at the same time and in proximity to other Grindr users—sometimes during the workday and sometimes while in transit between government facilities. For other Grindr users, he could infer where they lived, see where they traveled, and even guess at whom they were dating.

Intelligence agencies have a long and unfortunate history of trying to root out LGBTQ Americans from their workforce, but this wasn't Yeagley's intent. He didn't want anyone to get in trouble. No disciplinary actions were taken

against any employee of the federal government based on Yeagley's presentation. His aim was to show that buried in the seemingly innocuous technical data that comes off every cell phone in the world is a rich story—one that people might prefer to keep quiet. Or at the very least, not broadcast to the whole world. And that each of these intelligence and national security agencies had employees who were recklessly, if obliviously, broadcasting intimate details of their lives to anyone who knew where to look.

It wasn't only national security employees who could be compromised by this breach of privacy. Some Grindr users were not out to friends, family members, or their employers about their identity. As Yeagley showed, all that information was available for sale, for cheap. And it wasn't just Grindr, but rather any app that had access to a user's precise location—other dating apps, weather apps, games. Yeagley chose Grindr because it happened to generate a particularly rich set of data and its user base might be uniquely vulnerable. A Chinese company had obtained a majority stake in Grindr beginning in 2016—amping up fears in Washington that the data could be misused by a geopolitical foe. Until 1995, gay men and women were banned from having security clearances, owing in part to a belief among government counterintelligence agencies that their identities might make them vulnerable to being leveraged by an adversary—a belief that persists today.[*1]

Yeagley's point in these presentations was simple: data that most consumers didn't think twice about could be a resource for intelligence gathering and a threat to the privacy of citizens and the security of the United States. Either way, it needed to be guarded.[*2]

—

Having spent a decade in Washington as a reporter first for Politico and later at The Wall Street Journal and the Allbritton Journalism Institute, I'm used to coaxing stories out of sources and receiving tips. But the most extraordinary tale I've encountered fell into my lap during a wine-soaked dinner in the winter of 2018. Thanks to my dining companions that evening, I was given a chance glimpse inside a hidden world that I hadn't even begun to

contemplate. At the dinner, I was told of the existence of a government-linked effort to collect all the bits and bytes of advertising data that we were generating as consumers. It was being done through obscure contractors i the D.C. area, funded by the federal government. And because the U.S. governmentwas buying this data from a commercial provider, it was sanctioned by the law. At the time, that data was being used abroad in t global war on terror. But as you will come to understand, things that start abroad rarely stay there, and it wouldn't be long until this surveillance program came to America's shores.

As a demonstration, I was told to pick up my iPhone and select "Settings." From there, navigate to "Privacy," then "Advertising." There I found a toggle bar that asked if I wanted to "limit ad tracking." Below it, Apple explained that if I toggled on "Limit Ad Tracking," I would "opt out of receiving ads targeted to your interests."[*3]

It was that simple, I was told. This seemingly mundane setting on billions of mobile phones was a tangible clue that spoke to an entirely new kind of surveillanceprogram—onedesignedto track everyone.Everyone who possesses an iPhone or Android phone had all been given an "anonymized" advertising ID by Apple and Google, my companion explained. That number would be used to track our real-world movement, our internet browsing behavior, the apps we put on our phone, and much more. Billions of dollars had been poured into this system by America's largest corporations. And repository of data that rich and that detailed had attracted serious attention from the world's governments, which were opening their wallets to buy up information on everyone, rather than hacking it or getting it from secret court orders.

What I was learning over dinner was different from what Edward Snowden had revealed in 2013: that the U.S. government was running a massive surveillance apparatus with the cooperation of the largest tech an telecom companies,overseenby a court that operatesin secret. That surveillance effort was mostly focused on targets abroad, though some aspects of it touched on Americans and their data. Instead, what I was now learning about was a wholly separate effort. The government was buying its way into a

commercial marketplace, one that few consumers even knew existed. The little “Limit Ad Tracking” button was a way to limit some of the data that flowed into that marketplace. But no one can fully escape its clutches.

It would take more than five years after that dinner for me to fully understand the byzantine online ad ecosystem. Here’s how it works. Imagine a woman named Marcela who lives in the Philadelphia suburbs. She has a Google Pixel phone with the Weather Channel app installed. As she heads out the door to go on a jog, she sees overcast skies. So Marcela opens the app to check if the forecast calls for rain. By clicking on the Weather Channel’s bright blue icon, Marcela triggers a frenzy of digital activity all aimed at serving her a personalized ad. The Weather Channel has partnered with an entity called an advertising exchange to help pay for the app and deliver display ads to its millions of users. That exchange is basically a massive marketplace where billions of mobile devices and computers are telling a centralized server that they have an open ad space. And so in less than the blink of an eye after she opened the Weather Channel app, this machine goes to work. To deliver her the most relevant advertising, Marcela’s Google-assigned ad ID—called an AAID on Android phones—shared with the ad exchange so that it could serve her the most relevant possible ads based on what advertisers have inferred about her.

To the layperson, her AAID is a string of gibberish, something like bdca712j-fb3c-33ad-2324-0794d394m912. But to advertisers, it’s a gold mine. They know that bdca712j-fb3c-33ad-2324-0794d394m912 owns a Google Pixel device with the Nike Run Club app. They know that bdca712j-fb3c-33ad-2324-0794d394m912 often frequents runnersworld.com. And they know that bdca712j-fb3c-33ad-2324-0794d394m912 was lusting after a pair of new Vaporfly racing shoes. They know this because Nike, runnersworld.com, and Google are all plugged into the same advertising ecosystem, all aimed at understanding what consumers are interested in.

Advertisers use that information as they shape and deploy their ads. Say both Nike and Brooks, another running shoe brand, are trying to reach female running aficionados in a certain income bracket. Based on the huge amounts of data sloshing around, they might build an “audience”—essentially a huge

list of ad IDs of customers known or suspected to be in the market for running shoes. And they tell a digital ad exchange how much they're willing to pay to reach those consumers every time they load an app or a web page.

When Marcela loads the Weather Channel app, she sends reams of data back to the ad exchange. That includes the IP address of the phone, the type of phone and the operating system it's running, the carrier, the app in use, and the precise GPS coordinates of the phone. The exchange gets an array o technical data about how the phone is configured: what languages the browser is using, what version of the operating system is running, even what the screen resolution is set to. And finally, advertisers also get that pseudonymized advertising ID number. Technically, we can reset this number, but few people bother to. Few people even know they have one.

Users do have some control over what they share. And the advertisers have access only to whatever data the consumer grants them. If consume don't allow the app they're using to access GPS, the ad exchange can't pull the phone's GPS location, for example. (Or at least they aren't supposed to; not all the apps follow the rules, and Apple and Google don't always review the software in their app stores all that closely.)

Ad exchange bidding platforms do minimal due diligence on the hundreds or even thousands of entities that have a presence on their servers. So even the losing bidders still have access to all the consumer data that came off the phone during the bid request. An entire business model has been built on this: siphoning data off the real-time bidding networks, packaging it up, and reselling it to help businesses understand consumer behavior.

Geolocation is the single most valuable piece of commercial data to come off those devices. Understanding the movement of phones is now a multibillion-dollar industry. It can be used to deliver targeted advertising based on location for, say, a restaurant chain that wants to deliver targeted ads to people nearby. It can be used to measure consumer behavior and the effectiveness of advertising. How many people saw an ad and later visited a store? And the analytics can be used for planning and investment decisions. Where is the best location to put a new store? Will there be enough foot traffic to sustain such a business? Is the number of people visiting a certain

retailer going up or down this month, and what does that mean for the retailer's stock price?

But this kind of data is good for something else. It has remarkable surveillance potential. Why? Because what we do in the world with our devices cannot truly be anonymized. The fact that advertisers know Marcela as bdca712j-fb3c-33ad-2324-0794d394m912 as they're watching her move around the online and offline worlds offers her almost no privacy protection. Her habits and routines are unique to her. Our real-world movement is highly specific and personal to all of us. For many years, I lived in a small thirteen-unit walk-up in Washington, D.C. I was the only person waking up every morning at that address and going to the Journal's offices. Even if I was just an anonymized number, my behavior was as unique as a fingerprint even in a sea of hundreds of millions of others. There was no way to anonymize my identity in a data set like geolocation. Where a phone spends most of its evenings is a good proxy for where its owner lives. Advertisers know this Governments know this too. The only people it hasn't been explained to in a clear and resonant way is the general public.

Marcela—and the rest of us—were being tracked through a strange unholy alliance of big government and big business. Her data was being bought, sold, and traded in a marketplace that she didn't even know existed. The buyers were the largest advertisers and the biggest intelligence agencies.

This was the tantalizing and harrowing story laid out for me over that long-ago dinner. With growing horror as I swilled my wine and listened, I toggled the "Limit Ad Tracking" switch to on. And I did the only thing I knew how to do. I started reporting.

—

Writing this book took five years of my life, a lawsuit against the U.S. governmentunder the Freedom of Information Act, and hundredsof interviews with skittish and reluctant sources.What I have come to understand is this: the technology embedded in our phones, our computers, our cars, and our homes is part of a vast ecosystem of data collection a

analysis primarily aimed at understanding and in some cases manipulating our consumer behavior. Digital advertising is only one piece of it. Our public spaces are blanketed by networked cameras and other surveillance system put up in the name of public safety or personal security. And pretty much everything that emits a wireless signal of any kind—and today that list has grown to include routers, security cameras, televisions, home entertainment systems, Bluetooth keyboards, wireless headphones, and every single tire of every car manufactured since the mid-2000s, to name a few—can be and often is being covertly monitored. And the internet itself is built upon the backbone of a Cold War–era Defense Department computer network—with the routers, switches, packets, domain name lookups, and web addresses all subject to monitoring and manipulation in various ways.

Governments around the world—chief among them, the United States and its geopolitical rivals such as China and Russia—have accordingly come to view the internet not as a tool for self-expression, education, or commerce but as a mechanism for turning every single piece of consumer hardware and software on earth into a tool for intelligence gathering, "situational awareness," and in some cases social control. This is made possible not by hacking and breaking in—though governments do a lot of that too on their hardest and most valuable targets—nor by using expensive military hardware like overhead drones, spy planes, or satellites. In most cases, there's no need to go to such lengths because the consumer technology we use every day generates an unimaginable amount of data. And much of that data is for sale in opaque marketplaces and digital bazaars where the personal information of billions of consumers is bought, sold, and traded by the petabyte. Each piece of data on its own is not particularly valuable—a cell phone GPS ping here, the tire pressure reading on a car there. But woven together by government entities that operate in the shadows with multibillion-dollar budgets and powerful computer systems unavailable to the general public, the end result has been to blanket the globe in sensors, microphones, cameras, and scanners that are impossible to escape.

The modern digital ecosystem would not exist without surveillance—what the author and Harvard professor Shoshana Zuboff termed "surveillance

capitalism" in her landmark 2019 book, The Age of Surveillance Capitalism. Consumers, for better or worse, have begrudgingly come to accept that basic bargain. While we are not clear on the details, most of us have come to a rough understanding that our attention, behavior, and personal lives are being mined for behavioral insight by companies in exchange for free or discounted services. Zuboff recognized the government surveillance potential of the stores of data collected by the nation's largest corporations—data that she said was "raw material" for the system of surveillance capitalism she described. This book aims to fill in the details and bring into vivid relief just how deep the relationship between government intelligence agencies and our data goes.

This torrent of information is transforming government's relationship to its citizens. In some cases, it's for the better. Public health, city planning, transportation, medicine, and energy efficiency are being altered by insights unleashed using big data. But this revolution is also challenging every aspect of intelligence, law enforcement, and military operations—with profound consequences for the privacy, liberty, and dignity of citizens. Even in democratic countries, these activities are being done with scant public debate and little oversight from legislatures that barely understand the issues or the technology. And consumers and citizens have been kept in the dark—first by corporations, which do not want public scrutiny of the amount of data collection that occurs, and then by governments, which do not want to lose the specialized warrantless tracking capabilities that they have come to rely on.

Government lawyers have invoked the fact that this data is available publicly as the legal justification for its bulk acquisition and use. How much do we really care about our privacy if we've given this information away freely to the world's largest corporations? their argument goes. We all have consented to sharing intimate details of our lives for convenience and free services, and the government's counterterrorism and national security mission is much more important than selling patio furniture.

This is the paradox at the heart of this story. Corporations are loath to talk about the scale and scope of data collection because consumers find it

distasteful and there is money to be made. And governments have withheld critical details from us about how that data is being used in an expanding system of mass global surveillance while claiming that we have consented to its collection. The truth is that no consumer or citizen can know what data is being collected about them or how it's used, let alone consent. To say th anyone has consented to live in this world is a lie, because there is no way for the average consumer to even begin to understand the flow of data from their consumer technologies to corporate America and then to the security services of nearly every powerful nation on earth.

All of this has also been accomplished with almost no public discussion of what kind of world we're building in the twenty-first century.

—

This book is a chronicle of how different kinds of data became available for purchase by the U.S. government after 9/11 and the consequences for ou privacy. I've spent years trying to unravel this world—a fun house of mirrors draped in nondisclosure agreements, corporate trade secrets, needlessly classified contracts, misleading denials, and in some cases outright lies. This story is in rough chronological order, but because it sprawls across half a dozen government agencies over a two-decade time frame, that's not always possible and the story is not always perfectly linear.

Over years of thinking and reporting on this topic, I came to classify the data brokers in this story as belonging to one of four overlapping generations. First, there are consumer data brokers like Acxiom, Thomson Reuters, LexisNexis, and TransUnion that collect information like names, address histories, and consumer preferences. Second, there are social data providers that emerged to monitor the conversation on social media. Third, there ar advertising and location data brokers that sprang up to understand the movement of phones and the behavioral preferences of their owners. Finally, there are what might be called gray data providers that specialize in the most niche data sets.

My classification system is a vast oversimplification; the industry is in constant flux and large brokers like Thomson Reuters amass huge numbers of disparate data sets under one roof now. But these four types are a good way to think about the evolution of the industry. And as the market evolved, the government at each turn moved to take advantage of the possibilities offered by each new iteration of the data industry.

The four parts of this book roughly correspond to these four generations of data providers and the corresponding government efforts to capitalize on them. Part I traces the origins of consumer data brokers and the discover that after 9/11 they might have something to contribute to the counterterrorism mission. Part II documents the rise of social media and the government's early attempts to responsibly monitor it. Part III is about advertisingdata and smartphonesand the new vectors they offered to understand geographic behavior. And part IV is about the increasingly weird world of esoteric data that, without even knowing it, we're all generating, with vast consequences for our ability to move around the world without being subjected to persistent surveillance.

While this story focuses primarily on the activities of the U.S. government, the privacy issues raised in this book are global in scope. Every government on earth is eager to acquire data in any way possible to help it better understand the world.

—

There was never a grand overarching plan or conspiracy behind any of what I've described here. Rather, it's a story of different people at different periods in time working for different government agencies or contractors coming to the same realization: that data is available for sale and that it can be used for whatever mission is important at the time. This is a story about how a series of tiny, experimental programs, data vendors, and obscure contractors have brought us to the precipice of a digital panopticon—one built by corporate America and blessed by government lawyers.

"We are backing ourselves into a surveillance state," one former senior national security official told me one day in 2020. This was a man who had worked at the highest levels of American government and who had been intimately involved in the government's secret surveillance efforts after 9/11. But the growing aggregation of unclassified data gnawed at him far more than any secret surveillance program.

Information is power, he said. In the context of state power, data collection tilts the power toward the government and away from its citizens. What was being done today was arguably lawful but not thoughtful, he said—a myopic conversation among insiders that has excluded the general publi and failed to recognize a legal problem that was becoming a threat to the civil liberties and constitutional rights of Americans.

"Nobody should want this," he said.

S N

*1 It remains gospel in the national security community that being gay could pose a national security risk, especially if the employee is not open about their sexual orientation. The author James Kirchick argues that in reality LGBTQ Americans were unlikely to be shamed into betraying their country to hide their sexual identity. His book Secret City about the history of gay Washington makes the point that fears about compromise or blackmail were frequently used to drum LGBTQ Americans, usually gay men, out of official positions. "The belief was that because this was so terrible…the homosexual would go to any lengths to keep his secret a secret, and if that meant betraying his country…he would do it," Kirchick said in an interview. Kirchick argues there is not a single example in the entire espionage literature of someone being leveraged to betray their country because of their sexuality, pointing to a Defense Department study of more than a hundred international cases of espionage.

*2 Grindr has said in the years since Yeagley's demonstration it has drastically reduced the amount of data available to advertising exchanges and limited the number of data partnerships it has. It also doesn't sell ads in certain countries where being gay is a crime.

*3 Apple has made major changes in its settings since this time. As of this writing, "Limit Ad Tracking" is no longer buried deep in the iPhone settings. Instead, a very prominent box now asks users if they want to "allow apps to request to track." Disabling tracking is also prominently displayed in the privacy menu and has been relabeled "Tracking." As a result, there has been a drastic decrease in the number of iPhones being tracked in the years since

Apple made these changes, because users have opted out. As such, some of the techniques described in this book have been degraded, at least on iPhones.