the digital person

TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE

daniel j. solove

the digital person

Ex Machina: Law, Technology, and Society

General Editors: Jack M. Balkin *and* Beth Simone Noveck

The Digital Person

Technology and Privacy in the Information Age

Daniel J. Solove

the digital person

Technology and Privacy in the Information Age

daniel j. solove

NEW YORK UNIVERSITY PRESS *New York and London*

new york university press
New York and London
www.nyupress.org

Library of Congress Cataloging-in-Publication Data
Solove, Daniel J., 1972–
The digital person :
technology and privacy in the information age / Daniel J. Solove.
p. cm.—(Ex machina)
Includes bibliographical references and index.
ISBN 0–8147–9846–2 (cloth : alk. paper)
1. Data protection—Law and legislation—United States.
2. Electronic records—Access control—United States.
3. Public records—Law and legislation—United States.
4. Government information—United States.
5. Privacy, Right of—United States. I. Title. II. Series.
KF1263.C65S668 2004
343.7308'58—dc22 2004010188

New York University Press books are printed on acid-free paper,
and their binding materials are chosen for strength and durability.

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

In loving memory of

my grandma,

Jean

Contents

Acknowledgments

It is often said that books are written in solitude, but that wasn't true for this one. The ideas in this book were created in conversation with many wise friends and mentors. I owe them immense gratitude. Michael Sullivan has had an enormous influence on my thinking, and he has continually challenged me to strengthen my philosophical positions. Paul Schwartz has provided countless insights, and his work is foundational for the understanding of privacy law. Both Michael's and Paul's comments on the manuscript have been indispensable. I also must thank Judge Guido Calabresi, Naomi Lebowitz, Judge Stanley Sporkin, and Richard Weisberg, who have had a lasting impact on the way I think about law, literature, and life.

Charlie Sullivan deserves special thanks, although he disagrees with most of what I argue in this book. He has constantly forced me to better articulate and develop my positions. I may never convince him, but this book is much stronger for making the attempt.

So many other people are deserving of special mention, and if I were to thank them all to the extent they deserve, I would more than double the length of this book. Although I only list their names, my gratitude extends much further: Anita Allen, Jack Balkin, Carl Coleman, Howard Erichson, Timothy Glynn, Rachel Godsil, Eric Goldman, Chris Hoofnagle, Ted Janger, Jerry Kang, Orin Kerr, Raymond Ku, Erik Lillquist, Michael Risinger, Marc Rotenberg, Richard St. John, Chris Slobogin, Richard Sobel, Peter Swire, Elliot Turrini, and Benno Weisberg.

I greatly benefited from the comments I received when presenting my ideas, as well as portions of the manuscript, at conferences and symposia at Berkeley Law School, Cornell University, Emory Law School, Minnesota Law School, Seton Hall Law School, Stanford Law School, and Yale Law School.

My research assistants Peter Choy, Romana Kaleem, John Spaccarotella, and Eli Weiss provided excellent assistance throughout the writing of this book. Dean Pat Hobbs and Associate Dean Kathleen Boozang of Seton Hall Law School gave me generous support.

Don Gastwirth, my agent, shepherded me through the book publishing process with great enthusiasm and acumen. With unceasing attention, constant encouragement, and superb advice, he helped me find the perfect publisher. Deborah Gershenowitz at NYU Press believed in this project from the start and provided excellent editing.

Finally, I would like to thank my parents and grandparents. Their love, encouragement, and belief in me have made all the difference.

This book incorporates and builds upon some of my previously published work: *Privacy and Power: Computer Databases and Metaphors for Information Privacy,* 53 Stanford Law Review 1393 (2001); *Access and Aggregation: Privacy, Public Records, and the Constitution,* 86 Minnesota Law Review 1137 (2002); *Digital Dossiers and the Dissipation of Fourth Amendment Privacy,* 75 Southern California Law Review 1083 (2002); and *Identity Theft, Privacy, and the Architecture of Vulnerability,* 54 Hastings Law Journal 1227 (2003). These articles are really part of a larger argument, which I am delighted that I can now present in its entirety. The articles are thoroughly revised, and parts of different articles are now intermingled with each other. The argument can now fully unfold and develop. Privacy issues continue to change at a rapid pace, and even though these articles were written not too long ago, they were in need of updating. The arguments originally made in these articles have been strengthened by many subsequent discussions about the ideas I proposed. I have been forced to think about many issues more carefully and with more nuance. My understanding of privacy is a work in progress, and it has evolved since I began writing about it. This book merely represents another resting place, not the final word.

1 Introduction

We are in the midst of an information revolution, and we are only beginning to understand its implications. The past few decades have witnessed a dramatic transformation in the way we shop, bank, and go about our daily business—changes that have resulted in an unprecedented proliferation of records and data. Small details that were once captured in dim memories or fading scraps of paper are now preserved forever in the digital minds of computers, in vast databases with fertile fields of personal data. Our wallets are stuffed with ATM cards, calling cards, frequent shopper cards, and credit cards—all of which can be used to record where we are and what we do. Every day, rivulets of information stream into electric brains to be sifted, sorted, rearranged, and combined in hundreds of different ways. Digital technology enables the preservation of the minutia of our everyday comings and goings, of our likes and dislikes, of who we are and what we own. It is ever more possible to create an electronic collage that covers much of a person's life—a life captured in records, a digital person composed in the collective computer networks of the world.

We are currently confronting the rise of what I refer to as "digital dossiers." A dossier is a collection of detailed data about an individual. Dossiers are used in European courts to assemble information about a person in order to reach a judgment. Today, through the use of computers, dossiers are being constructed about all of us. Data is digitized into binary numerical form, which enables computers to store and manipulate it with unprecedented efficiency. There are hundreds of companies that are constructing gigantic databases of psychological profiles, amassing data about an individual's race, gender, income, hobbies, and purchases. Shards of data from our daily existence are now being assembled and analyzed—to investigate backgrounds, check credit, market products, and make a wide variety of decisions affecting our lives.

This book is about how we should understand and protect privacy in light of these profound technological developments. Our old conceptions of privacy are not up to the task. Much of the law pertaining to privacy is based on these conceptions, and as a result, it has failed to resolve the emerging privacy problems created by digital dossiers. This book aims to rethink longstanding notions of privacy to grapple with the consequences of living in an Information Age.

The Problems of Digital Dossiers

New Technologies and New Problems. In earlier times, communities were small and intimate. Personal information was preserved in the memories of friends, family, and neighbors, and it was spread by gossip and storytelling. Today, the predominant mode of spreading information is not through the flutter of gossiping tongues but through the language of electricity, where information pulses between massive record systems and databases. On the upside, this development means that individuals can more readily escape from the curious eyes of the community, freeing themselves from stifling social norms inhibiting individuality and creativity. On the downside, an ever-growing series of records is created about almost every facet of a person's life. As businesses and the government increasingly share personal information, digital dossiers about nearly every individual are being assembled. This raises serious concerns. The information gathered about us has become quite extensive, and it is being used in ways that profoundly affect our lives. Yet, we know little about how our personal information is being used, and we lack the power to do much about it.

Digital dossiers are constructed and used through three types of "information flow."[1] Information flow is a way of describing the movement of data. Like water in an elaborate system of plumbing, data flows through information pipelines linking various businesses, organizations, and government entities. First, information often flows between large computer databases of private-sector companies. Second, data flows from government public record systems to a variety of businesses in the private sector. Indeed, many companies construct their databases by culling personal data from public records. Third, information flows from the private sector to government agencies and law enforcement officials. The increase in digital dossiers has thus resulted in an elaborate lattice of information networking, where information is being stored, analyzed, and used in ways that have profound implications for society.

Even if we're not aware of it, the use of digital dossiers is shaping our lives. Companies use digital dossiers to determine how they do business with us; financial institutions use them to determine whether to give us credit; employers turn to them to examine our backgrounds when hiring; law enforcement officials draw on them to investigate us; and identity thieves tap into them to commit fraud.

Computer Databases. Computers and cyberspace have vastly increased our ability to collect, store, and analyze information. Today, it seems as if everyone is collecting information—the media, employers, businesses, and government. Countless companies maintain computerized records of their customers' preferences, purchases, and activities. There are hundreds of records detailing an individual's consumption. Credit card companies maintain information about one's credit card purchases. Video stores keep records about one's video rentals. Online retailers, such as Amazon.com, preserve records of all the books and other items a person buys. And there are hundreds of companies people aren't even aware of that maintain their personal information. For example, Wiland Services maintains a database of about 1,000 different points of information on over 215 million individuals.[2] Acxiom.com collects and sells data on consumers to marketers. In its InfoBase, it provides "[o]ver 50 demographic variables . . . including age, income, real property data, children's data, and others." It also contains data on education levels, occupation, height, weight, political affiliation, ethnicity, race, hobbies, and net worth.[3]

Computers enable marketers to collect detailed dossiers of personal information and to analyze it to predict the consumer's behavior. Through various analytic techniques, marketers construct models of what products particular customers will desire and how to encourage customers to consume. Companies know how we spend our money, what we do for a living, how much we earn, and where we live. They know about our ethnic backgrounds, religion, political views, and health problems. Not only do companies know what we have already purchased, but they also have a good idea about what books we will soon buy or what movies we will want to see.

Public Records. Imagine that the government had the power to compel individuals to reveal a vast amount of personal information about themselves—where they live; their phone numbers; their physical description; their photograph; their age; their medical problems; all of their legal transgressions throughout their lifetimes; the names of their parents, children, and spouses; their political party affiliations; where they work and what they do; the property that they own and its value; and sometimes even their psychotherapists' notes, doctors' records, and financial information.

Then imagine that the government routinely poured this information into the public domain—by posting it on the Internet where it could be accessed from all over the world, by giving it away to any individual or company that asked for it, or even by providing entire databases of personal information upon request. Think about how this information would be available to those who make important decisions about an individual's life and career—such as whether the individual will get a loan or a job. Also consider that, in many cases, the individual would not even know that the information is being used to make these decisions.

Imagine as well that this information would be traded among hundreds of private-sector companies that would combine it with a host of other information such as one's hobbies, purchases, magazines, organizations, credit history, and so on. This expanded profile would then be sold back to the government in order to investigate and monitor individuals more efficiently.

Stop imagining. What I described is a growing reality in the United States, and the threat posed to privacy is rapidly becoming worse. Federal, state, and local governments maintain public records spanning an individual's life from birth to death. These records contain a myriad of personal details. Until recently, public records were difficult to access—finding information about a person often involved a scavenger hunt through local offices to dig up records. But with the Internet, public records are increasingly being posted online, where anybody anywhere can easily obtain and search them.

Government Access. The data in digital dossiers increasingly flows from the private sector to the government, particularly for law enforcement use. Law enforcement agencies have long sought personal information about individuals from various companies and financial institutions to investigate fraud, white-collar crime, drug trafficking, computer crime, child pornography, and other types of criminal activity. In the aftermath of the terrorist attacks of September 11, 2001, the impetus for the government to gather personal information has greatly increased, since this data can be useful to track down terrorists and to profile airline passengers for more thorough searches. Detailed records of an individual's reading materials, purchases, diseases, and website activity enable the government to assemble a profile of an individual's finances, health, psychology, beliefs, politics, interests, and lifestyle. Many people communicate over the Internet using a screen name or pseudonym; the data in digital dossiers can unveil their identities as well as expose all of the people with whom they associate and do business.

The government has recently been exploring ways to develop technology to detect patterns of behavior based on our dossiers. In 2002, it was revealed that the Department of Defense was developing a program called Total Information Awareness (since renamed Terrorism Information Awareness). The program begins with the government amassing personal information from private-sector sources into a massive database of dossiers on individuals. Profiling technology is then used to detect those who are likely to be engaged in criminal activity. When Congress learned of Total Information Awareness, it halted the program because of its threat to privacy. However, the same type of collection and use of data envisioned by those who dreamed up Total Information Awareness is already being carried out by the government. The digital dossiers that continue to grow in the private sector and in public records are now becoming a tool for the government to monitor and investigate people.

What Is the Problem? The growing collection and use of personal information in digital form has long been viewed as problematic—a fear typically raised under the banner of "privacy." The use of personal information certainly presents privacy problems, but what exactly is the nature of these problems? Although the problems of personal information are understood as concerns over privacy, beyond this, they are often not well defined. How much weight should our vague apprehensions be given, especially considering the tremendous utility, profit, and efficiency of using personal information?

The answer to this question depends upon how these privacy problems are conceptualized. Unfortunately, so far, the problems have not been adequately articulated. Many discussions of privacy merely scratch the surface by simply pointing out a series of technological developments with the assumption that people will react with anxiety. Rarely do discussions about privacy delve deeper. We need a better understanding of the problems; we must learn how they developed, how they are connected, what precisely they threaten, and how they can be solved.

This book aims to reconceptualize privacy in today's world of rapidly changing technology. The task of conceptualizing the privacy problems digital dossiers create is of the utmost importance. As John Dewey aptly wrote, "a problem well put is half-solved."[4] We can't really solve a problem until we know the harm that it causes. A good diagnosis of a problem goes a long way toward finding solutions to it.

The goal of this book extends beyond articulating a new understanding of contemporary privacy problems; the book also aims to demonstrate the ways that the problems can be solved. In particular, this is a book about the law. A relatively robust amount of law has developed to protect privacy, but it has often failed to be effective when confronted by the problems of the Information Age. This book discusses why this has happened and what can be done about it.

Traditional Conceptions of Privacy

Traditionally, privacy violations have been understood in a particular manner. In this book, I contend that these ways of understanding privacy must be rethought in order to fully comprehend the problems with digital dossiers. This doesn't mean that these understandings are incorrect. They arose with earlier privacy problems and can certainly be of help in understanding digital dossiers. But these more traditional ways of understanding privacy don't account for key aspects of the unique problems the digital age has introduced.

Orwell's Big Brother. The dominant metaphor for modern invasions of privacy is Big Brother, the ruthless totalitarian government in George Orwell's novel *1984*. Big Brother oppresses its citizens, purges dissenters, and spies on everyone in their homes. The result is a cold, drab, grey world with hardly any space for love, joy, original thinking, spontaneity, or creativity. It is a society under total control. Although the metaphor has proven quite useful for a number of privacy problems, it only partially captures the problems of digital dossiers. Big Brother envisions a centralized authoritarian power that aims for absolute control, but the digital dossiers constructed by businesses aren't controlled by a central power, and their goal is not to oppress us but to get us to buy new products and services. Even our government is a far cry from Big Brother, for most government officials don't act out of malicious intent or a desire for domination. Moreover, Big Brother achieves its control by brutally punishing people for disobedience and making people fear they are constantly being watched. But businesses don't punish us so long as we keep on buying, and they don't make us feel as though we are being watched. To the contrary, they try to gather information as inconspicuously as possible. Making us feel threatened would undermine rather than advance the goal of unencumbered information collection. Finally, while Big Brother aims to control the most intimate details of a citizen's life, much of the information in digital dossiers is not intimate or unusual.

The Secrecy Paradigm. In another traditional way of understanding privacy that I refer to as the "secrecy paradigm," privacy is invaded by uncovering one's hidden world, by surveillance, and by the disclosure of concealed information. The harm such invasions cause consists of inhibition, self-censorship, embarrassment, and damage to one's reputation. The law is heavily influenced by this paradigm. As a result, if the information isn't secret, then courts often conclude that the information can't be private. However, this conception of privacy is not responsive to life in the modern Information Age, where most personal information exists in the record systems of hundreds of entities. Life today is fueled by information, and it is virtually impossible to live as an Information Age ghost, leaving no trail or residue.

The Invasion Conception. Under the traditional view, privacy is violated by the invasive actions of particular wrongdoers who cause direct injury to victims. Victims experience embarrassment, mental distress, or harm to their reputations. The law responds when a person's deepest secrets are exposed, reputation is tarnished, or home is invaded. This view, which I call the "invasion conception," understands privacy to be a kind of invasion, in which somebody invades and somebody is invaded. However, digital dossiers often do not result in any overt invasion. People frequently don't experience any direct injury when data about them is aggregated or transferred from one company to another. Moreover, many of the problems of digital dossiers emerge from the collaboration of a multitude of different actors with different purposes. Each step along the way is relatively small and innocuous, failing to cause harm that the invasion conception would recognize as substantial.

Rethinking Privacy

Digital dossiers pose significant problems, and for a more complete understanding of these issues, I turn to another metaphor—Franz Kafka's depiction of bureaucracy in *The Trial*. Kafka's novel chronicles the surreal nightmare of a person who is unexpectedly informed that he is under arrest but given no reason why. A bureaucratic court maintains a dossier about him, but he has no access to this information. Throughout the rest of the novel, the protagonist desperately attempts to find out why the Court is interested in his life, but his quest is hopeless—the Court is too clandestine and labyrinthine to be fully understood.

The Trial captures an individual's sense of helplessness, frustration, and vulnerability when a large bureaucratic organization has control over a vast dossier of details about one's life. Bureaucracy often results in a routinized and sometimes careless way of handling information—with little to no accountability. This makes people vulnerable to identity theft, stalking, and other harms. The problem is not simply a loss of control over personal information, nor is there a diabolical motive or plan for domination as with Big Brother. The problem is a bureaucratic process that is uncontrolled. These bureaucratic ways of using our information have palpable effects on our lives because people use our dossiers to make important decisions about us to which we are not always privy.

Thus far, the existing law protecting information privacy has not adequately responded to the emergence of digital dossiers. We need to better articulate what the problems are, what is at stake, and what precisely the law must do to solve the problems. We must rethink privacy for the Information Age.

A Roadmap for This Book

In part I, I explore the digital dossiers about individuals that are being assembled through computer databases and the Internet. I focus primarily on the activities of businesses. Chapter 2 traces the history of the developing privacy problems precipitated by computer databases and cyberspace. In chapter 3, I examine the prevailing ways that these privacy problems have been conceptualized. I discuss the predominance of the Orwell metaphor and why it must be supplemented with the Kafka metaphor. Chapter 4 discusses why the law of information privacy has failed to grapple adequately with the problem of digital dossiers. Chapter 5 responds to those who argue that the market (alone or with some minor tinkering) can appropriately deal with the problem. In chapter 6, I argue that beyond a set of individual rights, protecting privacy requires an architecture that regulates the way information may be collected and used. Consequently, protecting privacy must focus not merely on remedies and penalties for aggrieved individuals but on shaping an architecture to govern the ever-increasing data flows of the Information Age.

Part II turns to the way in which public records contribute to the problems of digital dossiers. Chapter 7 describes the increasing accumulation of personal information in public record systems and the emerging threats to privacy posed by the increased accessibility of this information as records are made available on the Internet. In chapter 8, I argue that the regulation of public records in the United States must be rethought in light of the new technologies in the Information Age, and I advance a theory about how to reconcile the tension between transparency and privacy. I also explore why regulating the access and use of public records will not infringe upon First Amendment rights.

Part III examines the problems created by the increasing government access to digital dossiers. In chapter 9, I explore in detail the numerous ways that the government is accessing personal information held by private-sector businesses and why this is problematic. Chapter 10 discusses how the Supreme Court has improperly interpreted the Fourth Amendment so that it doesn't apply to records maintained by third parties, a result that virtually prevents the Fourth Amendment from dealing with the problem of government access to digital dossiers. In the void left by the inapplicability of the Fourth Amendment, Congress has enacted a series of statutes to address the problem. As I explain, however, the statutes create a regulatory regime that is uneven, overly complex, and filled with gaps and loopholes. In chapter 11, I explore how the law should appropriately regulate government access to personal information maintained by the private sector.