"A fascinating journey into the almost surreal ways personal information is hoarded, used, and abused in the digital age." — *The Wall Street Journal*

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

2 The Rise of the Digital Dossier

We currently live in a world where extensive dossiers exist about each one of us. These dossiers are in digital format, stored in massive computer databases by a host of government agencies and private-sector companies. The problems caused by these developments are profound. But to understand the problems, we must first understand how they arose.

A History of Public-Sector Databases

Although personal records have been kept for centuries,[1] only in contemporary times has the practice become a serious concern. Prior to the nineteenth century, few public records were collected, and most of them were kept at a very local level, often by institutions associated with churches.[2] The federal government's early endeavors at collecting data consisted mainly in conducting the census. The first census in 1790 asked only four questions.[3] With each proceeding census, the government gathered more personal information. By 1860, 142 questions were asked.[4] When the 1890 census included questions about diseases, disabilities, and finances, it sparked a public outcry, ultimately leading to the passage in the early twentieth century of stricter laws protecting the confidentiality of census data.[5]

Government information collection flourished during the middle of the twentieth century. The creation and growth of government bureaucracy—spawning well over 100 federal agencies within the past century—led to an insatiable thirst for information about individuals. One such agency was the Social Security Administration, created in 1935, which assigned nine-digit numbers to each citizen and required extensive record-keeping of people's earnings.

Technology was a primary factor in the rise of information collection. The 1880 census required almost 1,500 clerks to tally information tediously by hand—and it took seven years to complete.[6] At the rapid rate of population growth, if a faster way could not be found to tabulate the information, the 1890 census wouldn't be completed before the 1900 census began. Fortunately, just in time for the 1890 census, a census official named Herman Hollerith developed an innovative tabulating device—a machine that read holes punched in cards.[7] Hollerith's new machine helped tabulate the 1890 census in under three years.[8] Hollerith left the Census Bureau and founded a small firm that produced punch card machines—a firm that through a series of mergers eventually formed the company that became IBM.[9]

IBM's subsequent rise to prosperity was due, in significant part, to the government's increasing need for data. The Social Security System and other New Deal programs required a vast increase in records that had to be kept about individuals. As a result, the government became one of the largest purchasers of IBM's punch card machines.[10] The Social Security Administration kept most of its records on punch cards, and by 1943 it had more than 100 million cards in storage.[11]

The advent of the mainframe computer in 1946 revolutionized information collection. The computer and magnetic tape enabled the systematic storage of data. As processing speeds accelerated and as memory ballooned, computers provided a vastly increased ability to collect, search, analyze, and transfer records.

Federal and state agencies began to computerize their records. The Census Bureau was one of the earliest purchasers of commercially available computers.[12] Social Security numbers (SSNs)—originally not to be used as identifiers beyond the Social Security System—became immensely useful for computer databases.[13] This is because SSNs en-

able data to be easily linked to particular individuals. In the 1970s, federal, state, and local governments—as well as the private sector—increasingly began to use them for identification.[14]

Beginning in the 1960s, the growing computerization of records generated a substantial buzz about privacy. Privacy captured the attention of the public, and a number of philosophers, legal scholars, and other commentators turned their attention to the threats to privacy caused by the rise of the computer.[15] Congress began to debate how to respond to these emerging developments.[16] In 1973, the U.S. Department of Health, Education, and Welfare (HEW) issued a report entitled *Records, Computers, and the Rights of Citizens,* which trenchantly articulated the growing concerns over computerized record systems:

> There was a time when information about an individual tended to be elicited in face-to-face contacts involving personal trust and a certain symmetry, or balance, between giver and receiver. Nowadays, an individual must increasingly give information about himself to large and relatively faceless institutions, for handling and use by strangers—unknown, unseen, and, all too frequently, unresponsive. Sometimes the individual does not even know that an organization maintains a record about him. Often he may not see it, much less contest its accuracy, control its dissemination, or challenge its use by others.[17]

These problems continued to escalate throughout the ensuing decades. Computers grew vastly more powerful, and computerized records became ubiquitous. The rise of the Internet in the 1990s added new dimensions to these problems, sparking a revolution in the collection, accessibility, and communication of personal data.

Today, federal agencies and departments maintain almost 2,000 databases,[18] including records pertaining to immigration, bankruptcy, licensing, welfare, and countless other matters. In a recent effort to track down parents who fail to pay child support, the federal government has created a vast database consisting of information about all people who obtain a new job anywhere in the nation. The database contains their SSNs, addresses, and wages.[19]

States maintain public records of arrests, births, criminal proceedings, marriages, divorces, property ownership, voter registration, workers' compensation, and scores of other types of records. State licensing regimes mandate that records be kept on numerous professionals such as doctors, lawyers, engineers, insurance agents, nurses, police, accountants, and teachers.

A History of Private-Sector Databases

Although the government played an integral role in the development of massive dossiers of personal information, especially early on, businesses soon began to play an even greater role. While the public-sector story concerns the quest for regulatory efficiency, the private-sector story involves money and marketing.

Long before the rise of nationwide advertising campaigns there was a personal relationship between merchant and customer. Local merchants lived next door to their customers and learned about their lives from their existence together in the community. To a large extent, marketing was done locally—by the peddler on the street or the shopkeeper on the corner. Mass marketing, which began in the nineteenth century and flourished in the twentieth century, transformed the nature of selling from personal one-to-one persuasion to large-scale advertising campaigns designed for the nameless, faceless American consumer.

Mass marketing consumed vast fortunes, and only a small fraction of the millions of people exposed to the ads would buy the products or services. Soon marketers discovered the power of a new form of marketing—targeted marketing. The idea was to figure out which people were most likely to consume a product and focus the advertising on them.

In the 1920s, the sales department of General Motors Corporation began an early experiment with targeted marketing. GM discovered that owners of Ford vehicles frequently didn't purchase a Ford as their next vehicle—so it targeted owners of two-year-old Fords and sent them a brochure on GM vehicles.[20] GM then began to send out questionnaires asking for consumer input into their products. GM be-

3 Kafka and Orwell

Reconceptualizing Information Privacy

The most widely discussed metaphor in the discourse of information privacy is George Orwell's depiction of Big Brother in *1984*. The use of the Big Brother metaphor to understand the database privacy problem is hardly surprising. Big Brother has long been the metaphor of choice to characterize privacy problems, and it has frequently been invoked when discussing police search tactics,[1] wiretapping and video surveillance,[2] and drug testing.[3] It is no surprise, then, that the burgeoning discourse on information privacy has seized upon this metaphor.

With regard to computer databases, however, Big Brother is incomplete as a way to understand the problem. Although the Big Brother metaphor certainly describes particular facets of the problem, it neglects many crucial dimensions. This oversight is far from inconsequential, for the way we conceptualize a problem has important ramifications for law and policy.

The Importance of Metaphor

A metaphor, as legal scholar Steven Winter aptly defines it, "is the imaginative capacity by which we relate one thing to another."[4] In

their groundbreaking analysis, linguistics professor George Lakoff and philosopher Mark Johnson observe that metaphors are not mere linguistic embellishments or decorative overlays on experience; they are part of our conceptual systems and affect the way we interpret our experiences.[5] Metaphor is not simply an act of description; it is a way of conceptualization. "The essence of metaphor," write Lakoff and Johnson, "is understanding and experiencing one kind of thing in terms of another."[6]

Much of our thinking about a problem involves the metaphors we use. According to legal philosopher Jack Balkin, "metaphoric models selectively describe a situation, and in so doing help to suppress alternative conceptions." Metaphors do not just distort reality but compose it; the "power [of metaphors] stems precisely from their ability to empower understanding by shaping and hence limiting it."[7]

Winter, as well as Lakoff and Johnson, focus on metaphors embodied in our thought processes, pervading the type of language we use.[8] The metaphors I speak of are not as deeply ingrained. Metaphors are tools of shared cultural understanding.[9] Privacy involves the type of society we are creating, and we often use metaphors to envision different possible worlds, ones that we want to live in and ones that we don't. Orwell's Big Brother is an example of this type of metaphor; it is a shared cultural narrative, one that people can readily comprehend and react to.

Ascribing metaphors is not only a descriptive endeavor but also an act of political theorizing with profound normative implications.[10] According to Judge Richard Posner, however, "it is a mistake to try to mine works of literature for political or economic significance" because works of literature are better treated as aesthetic works rather than "as works of moral or political philosophy."[11] To the contrary, literature supplies the metaphors by which we conceptualize certain problems, and Posner fails to acknowledge the role that metaphor plays in shaping our collective understanding. Metaphors function not to render a precise descriptive representation of the problem; rather, they capture our concerns over privacy in a way that is palpable, potent, and compelling. Metaphors are instructive not for their realism but for the way they direct our focus to certain social and political phenomena.

George Orwell's Big Brother

Orwell's Totalitarian World. Journalists, politicians, and jurists often describe the problem created by databases with the metaphor of Big Brother—the harrowing totalitarian government portrayed in George Orwell's *1984*.[12] Big Brother is an all-knowing, constantly vigilant government that regulates every aspect of one's existence. In every corner are posters of an enormous face, with "eyes [that] follow you about when you move" and the caption "BIG BROTHER IS WATCHING YOU."[13]

Big Brother demands complete obedience from its citizens and controls all aspects of their lives. It constructs the language, rewrites the history, purges its critics, indoctrinates the population, burns books, and obliterates all disagreeable relics from the past. Big Brother's goal is uniformity and complete discipline, and it attempts to police people to an unrelenting degree—even their innermost thoughts. Any trace of individualism is quickly suffocated.

This terrifying totalitarian state achieves its control by targeting the private life, employing various techniques of power to eliminate any sense of privacy. Big Brother views solitude as dangerous. Its techniques of power are predominantly methods of surveillance. Big Brother is constantly monitoring and spying; uniformed patrols linger on street corners; helicopters hover in the skies, poised to peer into windows. The primary surveillance tool is a device called a "telescreen" which is installed into each house and apartment. The telescreen is a bilateral television—individuals can watch it, but it also enables Big Brother to watch them:

> There was of course no way of knowing whether you were being watched at any given moment. How often, or on what system, the Thought Police plugged in on any individual wire was guesswork. It was even conceivable that they watched everybody all the time. . . . You had to live—did live, from habit that became instinct—in the assumption that every sound you made was overheard, and, except in darkness, every movement scrutinized.[14]

In *1984,* citizens have no way of discovering if and when they are being watched. This surveillance, both real and threatened, is

combined with swift and terrifying force: "People simply disappeared, always during the night. Your name was removed from the registers, every record of everything you had ever done was wiped out, your one-time existence was denied and then forgotten."[15]

Orwell's narrative brilliantly captures the horror of the world it depicts, and its images continue to be invoked in the legal discourse of privacy and information. "The ultimate horror in Orwell's imagined anti-utopia," observes sociologist Dennis Wrong, "is that men are deprived of the very capacity for cherishing private thoughts and feelings opposed to the regime, let alone acting on them."[16]

Panoptic Power. The telescreen functions similarly to the Panopticon, an architectural design for a prison, originally conceived by Jeremy Bentham in 1791.[17] In *Discipline and Punish,* Michel Foucault provides a compelling description of this artifice of power:

> [A]t the periphery, an annular building; at the centre, a tower; this tower is pierced with wide windows that open onto the inner side of the ring; the peripheric building is divided into cells, each of which extends the whole width of the building. . . . All that is needed, then, is to place a supervisor in a central tower and to shut up in each cell a madman, a patient, a condemned man, a worker or a schoolboy. By the effect of backlighting, one can observe from the tower, standing out precisely against the light, the small captive shadows in the cells of the periphery. They are like so many cages, so many small theatres, in which each actor is alone, perfectly individualized and constantly visible.[18]

The Panopticon is a device of discipline; its goal is to ensure order, to prevent plots and riots, to mandate total obedience. The Panopticon achieves its power through an ingenious technique of surveillance, one that is ruthlessly efficient. By setting up a central observation tower from which all prisoners can be observed and by concealing from them any indication of whether they are being watched at any given time, "surveillance is permanent in its effects, even if it is discontinuous in its action."[19] Instead of having hundreds of patrols and watchpersons, only a few people need to be in the tower. Those in the tower can watch any inmate but they cannot be

seen. By always being visible, by constantly living under the reality that one could be observed at any time, people assimilate the effects of surveillance into themselves. They obey not because they are monitored but because of their fear that they could be watched. This fear alone is sufficient to achieve control. The Panopticon is so efficient that nobody needs to be in the tower at all.

As Foucault observed, the Panopticon is not merely limited to the prison or to a specific architectural structure—it is a technology of power that can be used in many contexts and in a multitude of ways. In *1984,* the telescreen works in a similar way to the Panopticon, serving as a form of one-way surveillance that structures the behavior of those who are observed. The collection of information in cyberspace can be readily analogized to the telescreen. As we surf the Internet, information about us is being collected; we are being watched, but we do not know when or to what extent.

The metaphor of Big Brother understands privacy in terms of power, and it views privacy as an essential dimension of the political structure of society. Big Brother attempts to dominate the private life because it is the key to controlling an individual's entire existence: her thoughts, ideas, and actions.

The Ubiquity of the Metaphor. Big Brother dominates the discourse of information privacy. In 1974, when the use of computer databases was in its infancy, U.S. Supreme Court Justice William Douglas observed that we live in an Orwellian age in which the computer has become "the heart of a surveillance system that will turn society into a transparent world."[20] One state supreme court justice observed that the "acres of files" being assembled about us are leading to an "Orwellian society."[21]

Academics similarly characterize the problem.[22] In *The Culture of Surveillance,* sociologist William Staples observes that we have internalized Big Brother—we have created a Big Brother culture, where we all act as agents of surveillance and voyeurism.[23] "The specter of Big Brother has haunted computerization from the beginning," computer science professor Abbe Mowshowitz observes. "Computerized personal record-keeping systems, in the hands of police and intelligence agencies, clearly extend the surveillance capabilities of the state."[24]

Commentators have adapted the Big Brother metaphor to describe the threat to privacy caused by private-sector databases, often referring to businesses as "Little Brothers."[25] As sociologist David Lyon puts it: "Orwell's dystopic vision was dominated by the central state. He never guessed just how significant a decentralized consumerism might become for social control."[26] Legal scholar Katrin Byford writes: "Life in cyberspace, if left unregulated, thus promises to have distinct Orwellian overtones—with the notable difference that the primary threat to privacy comes not from government, but rather from the corporate world."[27] In *The End of Privacy,* political scientist Reg Whitaker also revises the Big Brother narrative into one of a multitude of Little Brothers.[28]

Internet "surveillance" can be readily compared to Orwell's telescreen. While people surf the web, companies are gathering information about them. As Paul Schwartz, a leading expert on privacy law, observes, the "Internet creates digital surveillance with nearly limitless data storage possibilities and efficient search possibilities." Instead of one Big Brother, today there are a "myriad" of "Big and Little Brothers" collecting personal data.[29]

Even when not directly invoking the metaphor, commentators frequently speak in its language, evoke its images and symbols, and define privacy problems in similar conceptual terms. Commentators view databases as having many of the same purposes (social control, suppression of individuality) and employing many of the same techniques (surveillance and monitoring) as Big Brother. David Flaherty, who served as the first Information and Privacy Commissioner for British Columbia, explains that the "storage of personal data can be used to limit opportunity and to encourage conformity." Dossiers of personal information "can have a limiting effect on behavior."[30] Oscar Gandy, a noted professor of communications and media studies, writes that "panopticism serves as a powerful metaphorical resource for representing the contemporary technology of segmentation and targeting."[31] As legal scholar Jerry Kang observes:

> [D]ata collection in cyberspace produces data that are detailed, computer-processable, indexed to the individual, and perma-

> nent. Combine this with the fact that cyberspace makes data collection and analysis exponentially cheaper than in real space, and we have what Roger Clarke has identified as the genuine threat of "dataveillance."[32]

Dataveillance, as information technology expert Roger Clarke defines it, refers to the "systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more persons."[33] According to political scientist Colin Bennet, "[t]he term *dataveillance* has been coined to describe the surveillance practices that the massive collection and storage of vast quantities of personal data have facilitated."[34] Dataveillance is thus a new form of surveillance, a method of watching not through the eye or the camera, but by collecting facts and data. Kang argues that surveillance is an attack on human dignity, interfering with free choice because it "leads to self-censorship."[35] Likewise, Paul Schwartz claims that data collection "creates a potential for suppressing a capacity for free choice: the more that is known about an individual, the easier it is to force his obedience."[36] According to this view, the problem with databases is that they are a form of surveillance that curtails individual freedom.

The Limits of the Metaphor. Despite the fact that the discourse appropriately conceptualizes privacy through metaphor and that the Big Brother metaphor has proven quite useful for a number of privacy problems, the metaphor has significant limitations for the database privacy problem. As illustrated by the history of record-keeping and databases in chapter 2, developments in record-keeping were not orchestrated according to a grand scheme but were largely ad hoc, arising as technology interacted with the demands of the growing public and private bureaucracies. Additionally, the goals of data collection have often been rather benign—or at least far less malignant than the aims of Big Brother. In fact, personal information has been collected and recorded for a panoply of purposes. The story of record-keeping and database production is, in the end, not a story about the progression toward a world ruled by Big Brother or a multitude of Little

Brothers. Instead, it is a story about a group of different actors with different purposes attempting to thrive in an increasingly information-based society.

The most significant shortcoming of the Big Brother metaphor is that it fails to focus on the appropriate form of power. The metaphor depicts a particular technique of power—surveillance. Certainly, monitoring is an aspect of information collection, and databases may eventually be used in ways that resemble the disciplinary regime of Big Brother. However, most of the existing practices associated with databases are quite different in character. Direct marketers wish to observe behavior so they can tailor goods and advertisements to individual differences. True, they desire consumers to act in a certain way (to purchase their product), but their limited attempts at control are far from the repressive regime of total control exercised by Big Brother. The goal of much data collection by marketers aims not at suppressing individuality but at studying it and exploiting it.

The most insidious aspect of the surveillance of Big Brother is missing in the context of databases: human judgment about the activities being observed (or the fear of that judgment). Surveillance leads to conformity, inhibition, and self-censorship in situations where it is likely to involve human judgment. Being observed by an insect on the wall is not invasive of privacy; rather, privacy is threatened by being subject to *human* observation, which involves judgments that can affect one's life and reputation. Since marketers generally are interested in aggregate data, they do not care about snooping into particular people's private lives. Much personal information is amassed and processed by computers; we are being watched not by other humans, but by machines, which gather information, compute profiles, and generate lists for mailing, emailing, or calling. This impersonality makes the surveillance less invasive.

While having one's actions monitored by computers does not involve immediate perception by a human consciousness, it still exposes people to the possibility of future review and disclosure. In the context of databases, however, this possibility is remote. Even when such data is used for marketing, marketers merely want to make a profit, not uproot a life or soil a reputation.

I do not, however, want to discount the dangerous effects of surveillance through the use of databases. Although the purposes of the users of personal data are generally not malignant, databases can still result in unintended harmful social effects. The mere knowledge that one's behavior is being monitored and recorded certainly can lead to self-censorship and inhibition. Foucault's analysis of surveillance points to a more subtle yet more pervasive effect: surveillance changes the entire landscape in which people act, leading toward an internalization of social norms that soon is not even perceived as repressive.[37] This view of the effects of surveillance raises important questions regarding the amount of normalization that is desirable in society. While our instincts may be to view all normalization as an insidious force, most theories of the good depend upon a significant degree of normalization to hold society together.

Although the effects of surveillance are certainly a part of the database problem, the heavy focus on surveillance miscomprehends the most central and pernicious effects of databases. Understanding the problem as surveillance fails to account for the majority of our activities in the world and web. A large portion of our personal information involves facts that we are not embarrassed about: our financial information, race, marital status, hobbies, occupation, and the like. Most people surf the web without wandering into its dark corners. The vast majority of the information collected about us concerns relatively innocuous details. The surveillance model does not explain why the recording of this non-taboo information poses a problem. The focus of the surveillance model is on the fringes—and often involves things we may indeed want to inhibit such as cult activity, terrorism, and child pornography.

Digital dossiers do cause a serious problem that is overlooked by the Big Brother metaphor, one that poses a threat not just to our freedom to explore the taboo, but to freedom in general. It is a problem that implicates the type of society we are becoming, the way we think, our place in the larger social order, and our ability to exercise meaningful control over our lives.

Franz Kafka's Trial

Kafka's Distopic Vision. Although we cannot arbitrarily adopt new metaphors, we certainly can exercise control over the metaphors we use. Since understanding our current society is an ongoing process, not a once-and-done activity, we are constantly in search of new metaphors to better comprehend our situation.

Franz Kafka's harrowing depiction of bureaucracy in *The Trial* captures dimensions of the digital dossier problem that the Big Brother metaphor does not.[38] *The Trial* opens with the protagonist, Joseph K., awakening one morning to find a group of officials in his apartment, who inform him that he is under arrest. K. is bewildered at why he has been placed under arrest: "I cannot recall the slightest offense that might be charged against me. But even that is of minor importance, the real question is, who accuses me? What authority is conducting these proceedings?" When he asks why the officials have come to arrest him, an official replies: "You are under arrest, certainly, more than that I do not know."[39] Instead of taking him away to a police station, the officials mysteriously leave.

Throughout the rest of the novel, Joseph K. begins a frustrating quest to discover why he has been arrested and how his case will be resolved. A vast bureaucratic court has apparently scrutinized his life and assembled a dossier on him. The Court is clandestine and mysterious, and court records are "inaccessible to the accused."[40] In an effort to learn about this Court and the proceedings against him, Joseph K. scuttles throughout the city, encountering a maze of lawyers, priests, and others, each revealing small scraps of knowledge about the workings of the Court. In a pivotal scene, Joseph K. meets a painter who gleaned much knowledge of the obscure workings of the Court while painting judicial portraits. The painter explains to K.:

> "The whole dossier continues to circulate, as the regular official routine demands, passing on to the highest Courts, being referred to the lower ones again, and then swinging backwards and forwards with greater or smaller oscillations, longer or shorter delays. . . . No document is ever lost, the Court never forgets anything. One day—quite unexpectedly—some Judge will take up

> the documents and look at them attentively. . . ." "And the case begins all over again?" asked K. almost incredulously. "Certainly" said the painter.[41]

Ironically, after the initial arrest, it is Joseph K. who takes the initiative in seeking out the Court. He is informed of an interrogation on Sunday, but only if he has no objection to it: "Nevertheless he was hurrying fast, so as if possible to arrive by nine o'clock, although he had not even been required to appear at any specific time."[42] Although the Court has barely imposed any authority, not even specifying when Joseph K. should arrive for his interrogation, he acts as if this Court operates with strict rules and makes every attempt to obey. After the interrogation, the Court seems to lose interest in him. Joseph K., however, becomes obsessed with his case. He wants to be recognized by the Court and to resolve his case; in fact, being ignored by the Court becomes a worse torment than being arrested.

As K. continues his search, he becomes increasingly perplexed by this unusual Court. The higher officials keep themselves hidden; the lawyers claim they have connections to Court officials but never offer any proof or results. Hardly anyone seems to have direct contact with the Court. In addition, its "proceedings were not only kept secret from the general public, but from the accused as well." Yet K. continues to seek an acquittal from a crime he hasn't been informed of and from an authority he cannot seem to find. As Joseph K. scurries through the bureaucratic labyrinth of the law, he can never make any progress toward his acquittal: "Progress had always been made, but the nature of the progress could never be divulged. The Advocate was always working away at the first plea, but it had never reached a conclusion."[43] In the end, Joseph K. is seized by two officials in the middle of the night and executed.

Kafka's *The Trial* best captures the scope, nature, and effects of the type of power relationship created by databases. My point is not that *The Trial* presents a more realistic descriptive account of the database problem than Big Brother. Like *1984, The Trial* presents a fictional portrait of a harrowing world, often exaggerating certain elements of society in a way that makes them humorous and absurd. Certainly, in the United States most people are not told that they are inexplicably

under arrest, and they do not expect to be executed unexpectedly one evening. *The Trial* is in part a satire, and what is important for the purposes of my argument are the insights the novel provides about society through its exaggerations. In the context of computer databases, Kafka's *The Trial* is the better focal point for the discourse than Big Brother. Kafka depicts an indifferent bureaucracy, where individuals are pawns, not knowing what is happening, having no say or ability to exercise meaningful control over the process. This lack of control allows the trial to completely take over Joseph K.'s life. *The Trial* captures the sense of helplessness, frustration, and vulnerability one experiences when a large bureaucratic organization has control over a vast dossier of details about one's life. At any time, something could happen to Joseph K.; decisions are made based on his data, and Joseph K. has no say, no knowledge, and no ability to fight back. He is completely at the mercy of the bureaucratic process.

As understood in light of the Kafka metaphor, the primary problem with databases stems from the way the bureaucratic process treats individuals and their information.

Bureaucracy. Generally, the term "bureaucracy" refers to large public and private organizations with hierarchical structures and a set of elaborate rules, routines, and processes.[44] I will use the term to refer not to specific institutions but to a particular set of practices—specifically, how bureaucratic processes affect and influence individuals subjected to them. Bureaucratic organization, sociologist Max Weber asserts, consists of a hierarchical chain-of-command, specialized offices to carry out particular functions, and a system of general rules to manage the organization.[45] Bureaucracy is not limited to government administration; it is also a feature of business management. The modern world requires the efficient flow of information in order to communicate, to deliver goods and services, to regulate, and to carry out basic functions. According to Weber, bureaucracy is "capable of attaining the highest degree of efficiency and is in this sense formally the most rational known means of exercising authority over human beings."[46] Bureaucratic processes are highly routinized, striving for increased efficiency, standardization of decisions, and the cultivation of specialization and expertise. As Paul Schwartz notes,

bureaucracy depends upon "vast quantities of information" that "relates to identifiable individuals."[47] Much of this information is important and necessary to the smooth functioning of bureaucracies.

Although bureaucratic organization is an essential and beneficial feature of modern society, bureaucracy also presents numerous problems. Weber observes that bureaucracy can become "dehumanized" by striving to eliminate "love, hatred, and all purely personal, irrational, and emotional elements which escape calculation."[48] Bureaucracy often cannot adequately attend to the needs of particular individuals—not because bureaucrats are malicious, but because they must act within strict time constraints, have limited training, and are frequently not able to respond to unusual situations in unique or creative ways. Schwartz contends that because bureaucracy does not adequately protect the dignity of the people it deals with, it can "weaken an individual's capacity for critical reflection and participation in society."[49] Additionally, decisions within public and private bureaucratic organizations are often hidden from public view, decreasing accountability. As Weber notes, "[b]ureaucratic administration always tends to exclude the public, to hide its knowledge and action from criticism as well as it can."[50] Bureaucratic organizations often have hidden pockets of discretion. At lower levels, discretion can enable abuses. Frequently, bureaucracies fail to train employees adequately and may employ subpar security measures over personal data. Bureaucracies are often careless in their uses and handling of personal information.

The problem with databases emerges from subjecting personal information to the bureaucratic process with little intelligent control or limitation, which results in our not having meaningful participation in decisions about our information. Bureaucratic decision-making processes are being exercised ever more frequently over a greater sphere of our lives, and we have little power or say within such a system, which tends to structure our participation along standardized ways that fail to enable us to achieve our goals, wants, and needs.

Bureaucracy and Power. The power effects of this relationship to bureaucracy are profound; however, they cannot adequately be explained by resorting only to the understanding of power in Orwell's

1984. Big Brother employs a coercive power that is designed to dominate and oppress. Power, however, is not merely prohibitive; as illustrated by Aldous Huxley in *Brave New World,* it composes our very lives and culture. Huxley describes a different form of totalitarian society—one controlled not by force, but by entertainment and pleasure. The population is addicted to a drug called Soma, which is administered by the government as a political tool to sedate the people. Huxley presents a narrative about a society controlled not by a despotic coercive government like Big Brother, but by manipulation and consumption, where people participate in their own enslavement. The government achieves obedience through social conditioning, propaganda, and other forms of indoctrination.[51] It does not use the crude coercive techniques of violence and force, but instead employs a more subtle scientific method of control—through genetic engineering, psychology, and drugs. Power works internally—the government actively molds the private life of its citizens, transforming it into a world of vapid pleasure, mindlessness, and numbness.

Despite the differences, power for both Orwell and Huxley operates as an insidious force employed for a particular design. *The Trial* depicts a different form of power. The power employed in *The Trial* has no apparent goal; any purpose remains shrouded in mystery. Nor is the power as direct and manipulative in design as that depicted by Orwell and Huxley. The Court system barely even cares about Joseph K. *The Trial* depicts a world that differs significantly from our traditional notions of a totalitarian state. Joseph K. was not arrested for his political views; nor did the Court manifest any plan to control people. Indeed, Joseph K. was searching for some reason why he was arrested, a reason that he never discovered. One frightening implication is that there was no reason, or if there were, it was absurd or arbitrary. Joseph K. was subjected to a more purposeless process than a trial. Indeed, the Court does not try to exercise much power over Joseph K. His arrest does not even involve his being taken into custody—merely a notification that he is under arrest—and after an initial proceeding, the Court makes no further effort even to contact Joseph K.

What is more discernible than any motive on the part of the Court or any overt exercise of power are the social effects of the power relationship between the bureaucracy and Joseph K. The power depicted

in *The Trial* is not so much a force as it is an element of relationships between individuals and society and government. These relationships have balances of power. What *The Trial* illustrates is that power is not merely exercised in totalitarian forms, and that relationships to bureaucracies which are unbalanced in power can have debilitating effects upon individuals—regardless of the bureaucracies' purposes (which may, in fact, turn out to be quite benign).

Under this view, the problem with databases and the practices currently associated with them is that they disempower people. They make people vulnerable by stripping them of control over their personal information. There is no diabolical motive or secret plan for domination; rather, there is a web of thoughtless decisions made by low-level bureaucrats, standardized policies, rigid routines, and a way of relating to individuals and their information that often becomes indifferent to their welfare.

The Interplay of the Metaphors. The Kafka and Orwell metaphors are not mutually exclusive. As I will discuss in more depth in part III of this book, the interplay of the metaphors captures the problems with government access to digital dossiers. In particular, the government is increasingly mining data from private-sector sources to profile individuals. Information about people is observed or recorded and then fed into computer programs that analyze the data looking for certain behavior patterns common to criminal or terrorist activity. This method of investigation and analysis employs secret algorithms to process information and calculate how "dangerous" or "criminal" a person might be. The results of these secret computations have palpable effects on people's lives. People can be denied the right to fly on an airplane without a reason or a hearing; or they can be detained indefinitely without the right to an attorney and without being told the reasons why.

In another example, political scientist John Gilliom's study of the surveillance of welfare recipients chronicles a world of constant observation coupled by an almost pathological bureaucracy.[52] Recipients must fill out mountains of paperwork, answer endless questions, and be routinely monitored. Often, they receive so little financial assistance that they resort to odd jobs to obtain more income, which, if

discovered, could make them ineligible for benefits. The system creates a strong incentive for transgression, severe penalties for any breach, and elaborate data systems that attempt to detect any malfeasance through automated investigations. The system combines pervasive surveillance with a bureaucratic process that has little compassion or flexibility.

A quote by noted playwright and author Friedrich Dürrenmatt best captures how surveillance and bureaucracy interrelate in the Information Age:

> [W]hat was even worse was the nature of those who observed and made a fool of him, namely a system of computers, for what he was observing was two cameras connected to two computers observed by two further computers and fed into computers connected to *those* computers in order to be scanned, converted, reconverted, and, after further processing by laboratory computers, developed, enlarged, viewed, and interpreted, by whom and where and whether at any point by human beings he couldn't tell.[53]

Surveillance generates information, which is often stored in record systems and used for new purposes. Being watched and inhibited in one's behavior is only one part of the problem; the other dimension is that the data is warehoused for unknown future uses. This is where Orwell meets Kafka.

Beyond the Secrecy Paradigm

Understanding the database privacy problem in terms of the Kafka metaphor illustrates that the problem with databases concerns the use of information, not merely keeping it secret. Traditionally, privacy problems have been understood as invasions into one's hidden world. Privacy is about concealment, and it is invaded by watching and by public disclosure of confidential information. I refer to this understanding of privacy as the "secrecy paradigm." This paradigm is so embedded in our privacy discourse that privacy is often represented visually by a roving eye, an open keyhole, or a person peeking through Venetian blinds.

Information about an individual, however, is often not secret, but is diffused in the minds of a multitude of people and scattered in various documents and computer files across the country. Few would be embarrassed by the disclosure of much of the material they read, the food they eat, or the products they purchase. Few would view their race, ethnicity, marital status, or religion as confidential. Of course, databases may contain the residue of scandals and skeletons—illicit websites, racy books, stigmatizing diseases—but since information in databases is rarely publicized, few reputations are tarnished. For the most part, the data is processed impersonally by computers without ever being viewed by the human eye. The secrecy paradigm focuses on breached confidentiality, harmed reputation, and unwanted publicity. But since these harms are not really the central problems of databases, privacy law often concludes that the information in databases is not private and is thus not entitled to protection. Indeed, one commentator defended DoubleClick's tracking of web browsing habits by stating:

> Over time, people will realize it's not Big Brother who's going to show up [at] your door in a black ski mask and take your kids away or dig deep into your medical history. This is a situation where you are essentially dropped into a bucket with 40 million people who look and feel a lot like you do to the advertising company.[54]

This commentator, viewing privacy with the Big Brother metaphor, focuses on the wrong types of harms and implicitly views only secret information as private.

The problem with databases pertains to the uses and practices associated with our information, not merely whether that information remains completely secret. Although disclosure can be a violation of privacy, this does not mean that avoiding disclosure is the sum and substance of our interest in privacy. What people want when they demand privacy with regard to their personal information is the ability to ensure that the information about them will be used only for the purposes they desire. Even regarding the confidentiality of information, the understanding of privacy as secrecy fails to recognize that individuals want to keep things private from some people but not

others. The fact that an employee criticizes her boss to a co-worker does not mean that she wants her boss to know what she said.

Helen Nissenbaum, a professor of information technology, is quite right to argue that we often expect privacy even when in public.[55] Not all activities are purely private in the sense that they occur in isolation and in hidden corners. When we talk in a restaurant, we do not expect to be listened to. A person may buy condoms or hemorrhoid medication in a store open to the public, but certainly expects these purchases to be private activities. Contrary to the notion that any information in public records cannot be private, there is a considerable loss of privacy by plucking inaccessible facts buried in some obscure document and broadcasting them to the world on the evening news. Privacy can be infringed even if no secrets are revealed and even if nobody is watching us.

The Aggregation Effect

The digital revolution has enabled information to be easily amassed and combined. Even information that is superficial or incomplete can be quite useful in obtaining more data about individuals. Information breeds information. For example, although one's SSN does not in and of itself reveal much about an individual, it provides access to one's financial information, educational records, medical records, and a whole host of other information. As law professor Julie Cohen notes, "[a] comprehensive collection of data about an individual is vastly more than the sum of its parts."[56] I refer to this phenomenon as the "aggregation effect." Similar to a Seurat painting, where a multitude of dots juxtaposed together form a picture, bits of information when aggregated paint a portrait of a person.

In the Information Age, personal data is being combined to create a digital biography about us. Information that appears innocuous can sometimes be the missing link, the critical detail in one's digital biography, or the key necessary to unlock other stores of personal information. But why should we be concerned about a biography that includes details about what type of soap a person buys, whether she prefers Pepsi to Coca-Cola, or whether she likes to shop at Macy's rather than Kmart? As legal scholar Stan Karas points out, the prod-

ucts we consume are expressive of our identities.[57] We have many choices in the products we buy, and even particular brands symbolize certain personality traits and personal characteristics. Karas notes that Pepsi has marketed itself to a younger, more rebellious consumer than Coca-Cola, which emphasizes old-fashioned and traditional images in its advertisements.[58] Whether punk, yuppie, or hippie, people often follow a particular consumption pattern that reflects the subculture with which they identify.[59]

Of course, the products we buy are not wholly reflective of our identities. A scene from Henry James's *Portrait of a Lady* best captures the complexities of the situation. Madame Merle, wise in the ways of the world yet jaded and selfish, is speaking to Isabel Archer, a young American lady in Europe full of great aspirations of living a bold and exceptional life, far beyond convention. Merle declares: "What shall we call our 'self'? Where does it begin? Where does it end? It overflows into everything that belongs to us—and then it flows back again. I know a large part of myself is the clothes I choose to wear. I've a great respect for *things*!" Isabel disagrees: "nothing that belongs to me is any measure of me." "My clothes only express the dressmaker," Isabel says, "but they don't express me. To begin with, it is not my own choice that I wear them; they've been imposed upon me by society."[60]

Merle is obsessed by things, and she views herself as deeply intertwined with her possessions. The objects she owns and purchases are deeply constitutive of her personality. Isabel, in her proud individualism, claims that she is vastly distinct from what she owns and wears. Indeed, for her, things are a tool for conformity; they do not express anything authentic about herself.

Yet Madame Merle has a point—the information is indeed expressive. But Isabel is right, too—this information is somewhat superficial, and it only partially captures who we are. Although the digital biography contains a host of details about a person, it captures a distorted persona, one who is constructed by a variety of external details.[61] Although the information marketers glean about us can be quite revealing, it still cannot penetrate into our thoughts and often only partially captures who we are.[62] Information about our property, our professions, our purchases, our finances, and our medical history does not tell the whole story. We are more than the bits of data we give

off as we go about our lives. Our digital biography is revealing of ourselves but in a rather standardized way. It consists of bits of information pre-defined based on the judgment of some entity about what categories of information are relevant or important. We are partially captured by details such as our age, race, gender, net worth, property owned, and so on, but only in a manner that standardizes us into types or categories. Indeed, database marketers frequently classify consumers into certain categories based on stereotypes about their values, lifestyle, and purchasing habits. As Julie Cohen observes, people are not simply "reducible to the sum of their transactions, genetic markers, and other measurable attributes."[63]

Our digital biography is thus an unauthorized one, only partially true and very reductive. We must all live with these unauthorized biographies about us, the complete contents of which we often do not get to see. Although a more extensive dossier might be less reductive in capturing our personalities, it would have greater controlling effects on an individual's life.

Not only are our digital biographies reductive, but they are often inaccurate. In today's bureaucratized world, one of the growing threats is that we will be subject to the inadvertence, carelessness, and mindlessness of bureaucracy. A scene from the darkly humorous movie *Brazil* illustrates this problem.[64] The movie opens with an exhausted bureaucrat swatting a fly, which inconspicuously drops into a typewriter, causes a jam, and results in him mistyping a letter in a person's name on a form. The form authorizes the arrest and interrogation of suspected rebels. In the next scene, an innocent man peacefully sits in his home with his family when suddenly scores of armor-clad police storm inside and haul him away.

These dangers are not merely the imaginary stuff of movies. The burgeoning use of databases of public record information by the private sector in screening job applicants and investigating existing employees demonstrates how errors can potentially destroy a person's career. For example, a Maryland woman wrongly arrested for a burglary was not cleared from the state's criminal databases. Her name and SSN also migrated to a Baltimore County database relating to child protective services cases. She was fired from her job as a substitute teacher, and only after she could establish that the information

was in error was she rehired. When she later left that job to run a day care center for the U.S. military, she was subjected to questioning about the erroneous arrest. Later on, when employed at as a child care director at a YMCA, she was terminated when her arrest record surfaced in a background clearance check. Since she could not have the error expunged in sufficient time, the job was given to another person. Only after several years was the error finally cleared from the public records.[65] As our digital biographies are increasingly relied upon to make important decisions, the problems that errors can cause will only escalate in frequency and magnitude.

To the extent that the digital biography is accurate, our lives are not only revealed and recorded, but also can be analyzed and investigated. Our digital biographies are being assembled by companies which are amassing personal information in public records along with other data. Collectively, millions of biographies can be searched, sorted, and analyzed in a matter of seconds. This enables automated investigations of individuals on a nationwide scale by both the government and the private sector. Increasingly, companies are conducting investigations which can have profound consequences on people's lives—such as their employment and financial condition. Employers are resorting to information brokers of public record information to assist in screening job applicants and existing employees. For example, the firm HireCheck serves over 4,000 employers to conduct background checks for new hires or current employees.[66] It conducts a national search of outstanding arrest warrants; a SSN search to locate the person's age, past and current employers, and former addresses; a driver record search; a search of worker's compensation claims "to avoid habitual claimants or to properly channel assignments"; a check of civil lawsuit records; and searches for many other types of information.[67] These investigations occur without any external oversight, and individuals often do not have an opportunity to challenge the results.

Forms of Dehumanization: Databases and the Kafka Metaphor

Expounding on the Kafka metaphor, certain uses of databases foster a state of powerlessness and vulnerability created by people's lack of

any meaningful form of participation in the collection and use of their personal information. Bureaucracy and power is certainly not a new problem. Databases do not cause the disempowering effects of bureaucracy; they exacerbate them—not merely by magnifying existing power imbalances but by transforming these relationships in profound ways that implicate our freedom. The problem is thus old and new, and its additional dimensions within the Information Age require extensive explication.

Impoverished Judgments. One of the great dangers of using information that we generally regard as private is that we often make judgments based on this private information about the person. As legal scholar Kenneth Karst warned in the 1960s, one danger of "a centralized, standardized data processing system" is that the facts stored about an individual "will become the only significant facts about the subject of the inquiry."[68] Legal scholar Jeffrey Rosen aptly observes, "Privacy protects us from being misdefined and judged out of context in a world of short attention spans, a world in which information can easily be confused with knowledge. True knowledge of another person is the culmination of a slow process of mutual revelation."[69]

Increased reliance upon the easily quantifiable and classifiable information available from databases is having profound social effects. The nature and volume of information affects the way people analyze, use, and react to information. Currently, we rely quite heavily on quantifiable data: statistics, polls, numbers, and figures. In the law alone, there is a trend to rank schools; to measure the influence of famous jurists by counting citations to their judicial opinions;[70] to assess the importance of law review articles by tabulating citations to them;[71] to rank law journals with an elaborate system of establishing point values for authors of articles;[72] and to determine the influence of academic movements by checking citations.[73] The goal of this use of empirical data is to eliminate the ambiguity and incommensurability of many aspects of life and try to categorize them into neat, tidy categories. The computer has exacerbated this tendency, for the increase in information and the way computers operate furthers this type of categorization and lack of judgment.[74] Indeed, in legal schol-

arship, much of this tendency is due to the advent of computer research databases, which can easily check for citations and specific terms.

In our increasingly bureaucratic and impersonal world, we are relying more heavily on records and profiles to assess reputation. As H. Jeff Smith, a professor of management and information technology, contends:

> [D]ecisions that were formerly based on judgment and human factors are instead decided according to prescribed formulas. In today's world, this response is often characterized by reliance on a rigid, unyielding process in which computerized information is given great weight. Facts that actually require substantial evaluation could instead be reduced to discrete entries in preassigned categories.[75]

Certainly, quantifiable information can be accurate and serve as the best way for making particular decisions. Even when quantifiable information is not exact, it is useful for making decisions because of administrative feasibility. Considering all the variables and a multitude of incommensurate factors might simply be impossible or too costly.

Nevertheless, the information in databases often fails to capture the texture of our lives. Rather than provide a nuanced portrait of our personalities, compilations of data capture the brute facts of what we do without the reasons. For example, a record of an arrest without the story or reason is misleading. The arrest could have been for civil disobedience in the 1960s—but it is still recorded as an arrest with some vague label, such as "disorderly conduct." It appears no differently from the arrest of a vandal. In short, we are reconstituted in databases as a digital person composed of data. The privacy problem stems paradoxically from the pervasiveness of this data—the fact that it encompasses much of our lives—as well as from its limitations—how it fails to capture us, how it distorts who we are.

Powerlessness and Lack of Participation. Privacy concerns an individual's power in the elaborate web of social relationships that encompasses

her life. Today, a significant number of these relationships involve interaction with public and private institutions. In addition to the myriad of public agencies that regulate the products we purchase, the environment, and the like, we depend upon private institutions such as telephone companies, utility companies, Internet service providers, cable service providers, and health insurance companies. We also depend upon companies that provide the products we believe are essential to our daily lives: hygiene, transportation, entertainment, news, and so on. Our lives are ensconced in these institutions, which have power over our day-to-day activities (through what we consume, read, and watch), our culture, politics, education, and economic well-being. We are engaged in relationships with these institutions, even if on the surface our interactions with them are as rudimentary and distant as signing up for services, paying bills, and requesting repairs. With many firms—such as credit reporting agencies—we do not even take affirmative steps to establish a relationship.

Companies are beginning to use personal information to identify what business experts call "angel" and "demon" customers.[76] Certain customers—the angels—are very profitable, but others—the demons—are not. Angel customers account for a large amount of a company's business whereas demon customers purchase only a small amount of goods and services and are likely to cost the company money. For example, a demon customer is one who uses up a company's resources by frequently calling customer service. Some business experts thus recommend that companies identify these types of customers through the use of personal information and treat them differently. For example, businesses might serve the angels first and leave the demons waiting; or they might offer the angels cheaper prices; or perhaps, they might even try to turn the demons away entirely.[77] The result of companies moving in this direction is that people will be treated differently and may never know why. Even before the concept of angel and demon customers was articulated, one bank routinely denied credit card applications from college students majoring in literature, history, and art, based on the assumption that they would not be able to repay their debts. The bank's practice remained a secret until the media ran a story about it.[78]

We are increasingly not being treated as equals in our relationships with many private-sector institutions. Things are done to us; decisions are made about us; and we are often completely excluded from the process. With considerably greater frequency, we are ending up frustrated with the outcome. For example, complaints about credit reporting agencies to the Federal Trade Commission have been rapidly escalating, with 8,000 in 2001 and over 14,000 in 2002.[79]

Privacy involves the ability to avoid the powerlessness of having others control information that can affect whether an individual gets a job, becomes licensed to practice in a profession, or obtains a critical loan. It involves the ability to avoid the collection and circulation of such powerful information in one's life without having any say in the process, without knowing who has what information, what purposes or motives those entities have, or what will be done with that information in the future. Privacy involves the power to refuse to be treated with bureaucratic indifference when one complains about errors or when one wants certain data expunged. It is not merely the collection of data that is the problem—it is our complete lack of control over the ways it is used or may be used in the future.

Problematic Information Gathering Techniques. This powerlessness is compounded by the fact that the process of information collection in America is clandestine, duplicitous, and unfair. The choices given to people over their information are hardly choices at all. People must relinquish personal data to gain employment, procure insurance, obtain a credit card, or otherwise participate like a normal citizen in today's economy. Consent is virtually meaningless in many contexts. When people give consent, they must often consent to a total surrender of control over their information.

Collection of information is often done by misleading the consumer. General Electric sent a supposedly anonymous survey to shareholders asking them to rate various aspects of the company. Unbeknownst to those surveyed, the survey's return envelope was coded so that the responses could be matched to names in the company's shareholder database.[80]

Some information is directly solicited via registration questionnaires or other means such as competitions and sweepstakes. The

warranty registration cards of many products—which ask a host of lifestyle questions—are often sent not to the company that makes the product but to National Demographics and Lifestyles Company at a Denver post office box. This company has compiled information on over 20 million people and markets it to other companies.[81] Often, there is an implicit misleading notion that consumers must fill out a registration questionnaire in order to be covered by the warranty.

Frequent shopper programs and discount cards—which involve filling out a questionnaire and then carrying a special card that provides discounts—enable the scanner data to be matched to data about individual consumers.[82] This technique involves offering savings in return for personal information and the ability to track a person's grocery purchases.[83] However, there are scant disclosures that such an exchange is taking place, and there are virtually no limits on the use of the data.

Conde Nast Publications Inc. (which publishes the *New Yorker, Vanity Fair, Vogue,* and other magazines) recently sent out a booklet of 700 questions asking detailed information about an individual's hobbies, shopping preferences, health (including medications used, acne problems, and vaginal/yeast infections), and much more. Almost 400,000 people responded. In return for the data, the survey said: "Just answer the questions below to start the conversation and become part of this select group of subscribers to whom marketers listen first." Conde Nast maintains a database of information on 15 million people. Stephen Jacoby, the vice president for marketing and databases, said: "What we're trying to do is enhance the relationship between the subscriber and their magazine. In a sense, it's a benefit to the subscriber."[84]

There is no "conversation" created by supplying the data. Conde Nast does not indicate how the information will be used. It basically tries to entice people to give information for a vague promise of little or no value. While the company insists that it will not share information with "outsiders," it does not explain who constitutes an "outsider." The information remains in the control of the company, with no limitations on use. Merely informing the consumer that data may be sold to others is an inadequate form of disclosure. The consumer

does not know how many times the data will be resold, to whom it will be sold, or for what purposes it will be used.

Irresponsibility and Carelessness. A person's lack of control is exacerbated by the often thoughtless and irresponsible ways that bureaucracies use personal information and their lack of accountability in using and protecting the data. In other words, the problem is not simply a lack of individual control over information, but a situation where *nobody* is exercising meaningful control over the information.

In bureaucratic settings, privacy policy tends to fall into drift and be reactionary. In a detailed study of organizations such as banks, insurance companies, and credit reporting agencies, H. Jeff Smith concluded that all of the organizations "exhibited a remarkably similar approach: the policy-making process, which occurred over time, was a wandering and reactive one." According to a senior executive at a health insurance company, "We've been lazy on the privacy [issues] for several years now, because we haven't had anybody beating us over the head about them." According to Smith, most executives in the survey were followers rather than leaders: "[M]ost executives wait until an external threat forces them to consider their privacy policies."[85]

Furthermore, there have been several highly publicized instances where companies violated their own privacy policies. Although promising its users that their information would remain confidential, the website GeoCities collected and sold information about children who played games on the site.[86] RealNetworks, Inc. secretly collected personal information about its users in direct violation of its privacy policy. And a website for young investors promised that the data it collected about people's finances would remain anonymous, but instead it was kept in identifiable form.[87]

More insidious than drifting and reactionary privacy policies are irresponsible and careless uses of personal information. For example, Metromail Corporation, a seller of direct marketing information, hired inmates to enter the information into databases. This came to light when an inmate began sending harassing letters that were sexually explicit and filled with intimate details of people's lives.[88] A television reporter once paid $277 to obtain from Metromail a list of over

5,000 children living in Pasadena, California. The reporter gave the name of a well-known child molester and murderer as the buyer.[89] These cases illustrate the lack of care and accountability by the corporations collecting the data.

McVeigh v. Cohen[90] best illustrates this problem. A highly decorated 17-year veteran of the Navy sought to enjoin the Navy from discharging him under the statutory policy known as "Don't Ask, Don't Tell, Don't Pursue."[91] When responding to a toy drive for the crew of his ship, Tim McVeigh (no relation to the Oklahoma City bomber) accidentally used the wrong email account, sending a message under the alias "boysrch." He signed the email "Tim" but included no other information. The person conducting the toy drive searched through the member profile directory of America Online (AOL), where she learned that "boysrch" was an AOL subscriber named Tim who lived in Hawaii and worked in the military. Under marital status, he had identified himself as "gay." The ship's legal adviser began to investigate, suspecting that "Tim" was McVeigh. Before speaking to McVeigh, and without a warrant, the legal adviser had a paralegal contact AOL for more information. The paralegal called AOL's toll-free customer service number and, without identifying himself as a Navy serviceman, concocted a story that he had received a fax from an AOL customer and wanted to confirm who it belonged to. Despite a policy of not giving out personal information, the AOL representative told him that the customer was McVeigh. As a result, the Navy sought to discharge McVeigh.

In *Remsburg v. Docusearch, Inc.*,[92] a man named Liam Youens began purchasing information about Amy Lynn Boyer from a company called Docusearch. He requested Boyer's SSN, and Docusearch obtained it from a credit reporting agency and provided it to him. Youens then requested Boyer's employment address, so Docusearch hired a subcontractor, who obtained it by making a "pretext" phone call to Boyer. By lying about her identity and the reason for the call, the subcontractor obtained the address from Boyer. Docusearch then gave the address to Youens, who went to Boyer's workplace and shot and killed her. Docusearch supplied the information without ever asking who Youens was or why he was seeking the information.

Within the past few years, explicit details of 90 psychotherapy patients' sex lives, as well as their names, addresses, telephone num-

bers, and credit card numbers, were inadvertently posted on the Internet.[93] A banker in Maryland who sat on a state's public health commission checked his list of bank loans with records of people with cancer in order to cancel the loans of the cancer sufferers.[94] A hacker illegally downloaded thousands of patients' medical files along with their SSNs from a university medical center.[95] Due to a mix-up, a retirement plan mailed financial statements to the wrong people at the same firm.[96] Extensive psychological records describing the conditions of over 60 children were inadvertently posted on the University of Montana's website.[97] An employee of a company obtained 30,000 credit reports from a credit reporting agency and peddled them to others for use in fraud and identity theft.[98] Health information and SSNs of military personnel and their families were stolen from a military contractor's database.[99]

In sum, the privacy problem created by the use of databases stems from an often careless and unconcerned bureaucratic process—one that has little judgment or accountability—and is driven by ends other than the protection of people's dignity. We are not just heading toward a world of Big Brother or one composed of Little Brothers, but also toward a more mindless process—of bureaucratic indifference, arbitrary errors, and dehumanization—a world that is beginning to resemble Kafka's vision in *The Trial*.

9 Government Information Gathering

Thus far, I have discussed how personal information is being more readily collected, stored, transferred, and combined with other information. Part I of this book discussed the problems of information flow among various businesses, and part II focused on information flows from the government to the private sector. But there is another problematic type of information flow that is rapidly escalating—data transfers from the private sector to the government. The vast digital dossiers being constructed by businesses are becoming an increasingly desirable resource for law enforcement officials. And this threatens to transform the relationship between government and citizen in some very troubling ways.

Third Party Records and the Government

Earlier in this book, I described the extensive amount of information that companies are stockpiling about us. To live in the modern world, we must enter into numerous relationships with other people and businesses: doctors, lawyers, businesses, merchants, magazines,

newspapers, banks, credit card companies, employers, landlords, ISPs, insurance companies, phone companies, and cable companies. The list goes on and on. Our relationships with all of these entities generate records containing personal information necessary to establish an account and record our transactions and preferences. We are becoming a society of records, and these records are not held by us, but by third parties.

These record systems are becoming increasingly useful to law enforcement officials. Personal information can help the government detect fraud, espionage, fugitives, drug distribution rings, and terrorist cells. Information about a person's financial transactions, purchases, and religious and political beliefs can assist the investigation of suspected criminals and can be used to profile people for more thorough searches at airports.

The government, therefore, has a strong desire to obtain personal information found in records maintained by third parties. For instance, from pen registers and trap and trace devices, the government can obtain a list of all the phone numbers dialed to or from a particular location, potentially revealing the people with whom a person associates. From bank records, which contain one's account activity and check writing, the government can discover the various companies and professionals that a person does business with (ISP, telephone company, credit card company, magazine companies, doctors, attorneys, and so on). Credit card company records can reveal where one eats and shops. The government can obtain one's travel destinations and activities from travel agent records. From hotel records, it can discover the numbers a person dialed and the pay-per-view movies a person watched.[1] The government can obtain one's thumbprint from car rental companies that collect them to investigate fraud.[2] From video stores, the government can access an inventory of the videos that a person has rented.

The government can also glean a wealth of information from the extensive records employers maintain about their employees.[3] Employers frequently monitor their employees.[4] Some use software to track how employees surf the Internet.[5] Employers often record information about an employee's email use, including back-up copies of the contents of email. A number of employers also conduct drug test-

ing, and many require prospective employees to answer questionnaires asking about drug use, finances, psychological treatment, marital history, and sexuality.[6] Some even require prospective hires to take a psychological screening test.[7]

Landlords are another fertile source of personal information. Landlord records often contain financial, employment, and pet information, in addition to any tenant complaints. Many landlords also maintain log books at the front desk where visitors sign in. Some apartment buildings use biometric identification devices, such as hand scanners, to control access to common areas such as gyms.

Increasingly, companies and entities that we have never established any contact with nevertheless have dossiers about us. Credit reporting agencies maintain information relating to financial transactions, debts, creditors, and checking accounts. The government can also find out details about people's race, income, opinions, political beliefs, health, lifestyle, and purchasing habits from the database companies that keep extensive personal information on millions of Americans.

Beyond the records described here, the Internet has the potential to become one of the government's greatest information gathering tools. There are two significant aspects of the Internet that make it such a revolutionary data collection device. First, it gives many individuals a false sense of privacy. The secrecy and anonymity of the Internet is often a mirage. Rarely are people truly anonymous because ISPs keep records of a subscriber's screen name and pseudonyms. ISP account information includes the subscriber's name, address, phone numbers, passwords, information about web surfing sessions and durations, and financial information.[8] By learning a person's screen name, the government can identify who posted messages in newsgroups or conversed in chatrooms.

At the government's request, an ISP can keep logs of the email addresses with which a person corresponds. Further, if a person stores email that is sent and received with the ISP, the government can obtain the contents of those emails.

Second, the Internet is unprecedented in the degree of information that can be gathered and stored. It is one of the most powerful generators of records in human history. As discussed in chapter 2, websites

often accumulate a great deal of information about their users, from transactional data to information collected through cookies. The government can glean a substantial amount of information about visitors to a particular website. From Internet retailers, the government can learn about the books, videos, music, and electronics that one purchases. Some Internet retailers, such as Amazon.com, record all the purchases a person has made throughout many years. Based on this information, the government can discover a consumer's interests, political views, religious beliefs, and lifestyle.

The government may also obtain information from websites that operate personalized home pages. Home pages enable users to keep track of the stocks they own, favorite television channels, airfares for favorite destinations, and news of interest. Other websites, such as Microsoft Network's calendar service, allow users to maintain their daily schedule and appointments. Further, as discussed in chapter 2, there are database companies that amass extensive profiles of people's websurfing habits.

While life in the Information Age has brought us a dizzying amount of information, it has also placed a profound amount of information about our lives in the hands of numerous entities. As discussed earlier, these digital dossiers are increasingly becoming digital biographies, a horde of aggregated bits of information combined to reveal a portrait of who we are based upon what we buy, the organizations we belong to, how we navigate the Internet, and which shows and videos we watch. This information is not held by trusted friends or family members, but by large bureaucracies that we do not know very well or sometimes do not even know at all.

Government–Private-Sector Information Flows

In late 2002, the news media reported that the Department of Defense was planning a project known as Total Information Awareness (TIA). The project was to be run by John Poindexter, who had been convicted in 1990 for his activities during the Iran-contra scandal. TIA envisioned the creation of a gigantic government database of personal information, including data culled from private-sector entities concerning finances, education, travel, health, and so on. This infor-

mation would then be analyzed under various models to detect patterns and profiles for terrorist activities.[9] The website for the project contained the symbol of a pyramid with beams of light emanating from an eye at the top. Next to the pyramid was a globe, illuminated by the light. Underneath the image were the words *scientia est potentia*—"knowledge is power."[10]

When TIA broke as a major news story, civil liberties groups and many commentators and politicians voiced stinging criticism. In a *New York Times* editorial, William Safire wrote that Poindexter "is determined to break down the wall between commercial snooping and secret government intrusion. . . . And he has been given a $200 million budget to create computer dossiers on 300 million Americans."[11] After these outcries, the pyramid and eye logo was quickly removed from the Department of Defense website. The Senate amended its spending bill in January 2003 to temporarily suspend funding for TIA until the details of the program were explained to Congress.[12] In May 2003, the Department of Defense issued its report to Congress, renaming the program "Terrorism Information Awareness" and declaring (without specifying how) that it would protect privacy. Later on, in July, the Senate voted unanimously to stop funding for TIA. The program had been killed.

But TIA is only one part of the story of government access to personal information and its creation of dossiers on American citizens. In fact, for quite some time, the government has been increasingly contracting with businesses to acquire databases of personal information. Database firms are willing to supply the information and the government is willing to pay for it. Currently, government agencies such as the FBI and IRS are purchasing databases of personal information from private-sector companies.[13] A private company called ChoicePoint, Inc. has amassed a database of 10 billion records and has contracts with at least 35 federal agencies to share the data with them. In 2000, the Justice Department signed an $8 million contract with ChoicePoint, and the IRS reached a deal with the company for between $8 and $12 million. ChoicePoint collects information from public records from around the country and then combines it with information from private detectives, the media, and credit reporting firms. This data is indexed by people's SSNs. The Center for Medicare

and Medicaid Services uses ChoicePoint's data to help it identify fraudulent Medicare claims by checking health care provider addresses against ChoicePoint's list of "high-risk and fraudulent business addresses." ChoicePoint's information is not only used by government agencies but also by private-sector employers to screen new hires or investigate existing employees.[14]

ChoicePoint's information is a mixture of fact and fiction. There are a number of errors in the records, such as when a ChoicePoint report falsely indicated that a woman was a convicted drug dealer and shoplifter, resulting in her being fired from her job.[15] ChoicePoint also had a hand in the 2000 presidential election problems in Florida. ChoicePoint supplied Florida officials with a list of 8,000 "ex-felons" to eliminate from their voter lists.[16] However, many of the 8,000 were not guilty of felonies, only misdemeanors, and were legally eligible to vote. Although the error was discovered prior to the election and officials tried to place the individuals back on the voter rolls, the error may have led to some eligible voters being turned away at the polls.

Additionally, many states have joined together to create a database system called Multi-State Anti-Terrorism Information Exchange, or MATRIX for short. Run by SeisInt, Inc., a private-sector company in Florida, MATRIX contains personal information gathered from public records and from businesses. In its vast fields of data, MATRIX includes people's criminal histories, photographs, property ownership, SSNs, addresses, bankruptcies, family members, and credit information. The federal government has provided $12 million to help support the program.[17]

A second form of information flow from the private sector to the government emerges when the government requests private-sector records for particular investigations or compels their disclosure by subpoena or court order. Voluntary disclosure of customer information is within the third party company's discretion. Further, whether a person is notified of the request and given the opportunity to challenge it in court is also within the company's discretion.

The September 11, 2001 terrorist attacks changed the climate for private sector-to-government information flows. Law enforcement officials have a greater desire to obtain information that could be helpful in identifying terrorists or their supporters, including infor-

mation about what people read, the people with whom they associate, their religion, and their lifestyle. Following the September 11 attack, the FBI simply requested records from businesses without a subpoena, warrant, or court order.[18] Recently, Attorney General John Ashcroft has revised longstanding guidelines for FBI surveillance practices. Under the previous version, the FBI could monitor public events and mine the Internet for information only when "facts or circumstances reasonably indicate that a federal crime has been, is being, or will be committed."[19] Under the revised version, the FBI can engage in these types of information gathering without any requirement that it be part of a legitimate investigation or related in any manner to criminal wrongdoing. The FBI can now collect "publicly available information, whether obtained directly or through services or resources (whether nonprofit or commercial) that compile or analyze such information; and information voluntarily provided by private entities."[20] Further, the FBI can "carry out general topical research, including conducting online searches and accessing online sites and forums."[21]

In conjunction with the government's greater desire for personal information, the private sector has become more willing to supply it. Background check companies, for instance, experienced a large boost in business after September 11.[22] Several large financial companies developed agreements to provide information to federal law enforcement agencies.[23] Indeed, in times of crisis or when serious crimes are at issue, the incentives to disclose information to the government are quite significant. Shortly after September 11, around 200 universities admitted to giving the FBI access to their records on foreign students—often without a subpoena or court order.[24] In violation of its privacy policy, JetBlue Airlines shared the personal data of 1 million customers with Torch Concepts, an Alabama company contracting with the Defense Department to profile passengers for security risks. Torch combined the JetBlue data with SSNs, employment information, and other details obtained from Acxiom, Inc., a database marketing company.[25] In a similar incident, Northwest Airlines secretly turned over to NASA its customer data—including addresses, phone numbers, and credit card information—for use in a government data mining project.[26] In a December 2002 survey of nearly 800

chief security officers, almost 25 percent said that they would supply information to the government without a court order, with 41 percent doing so in cases involving national security.[27]

When businesses refuse to cooperate, the government can compel production of the information by issuing a subpoena or obtaining a court order. These devices are very different from warrants because they offer little protection to the individual being investigated. Notification of the target of the investigation is often within the discretion of the third party. Further, it is up to the third party to challenge the subpoena. So, rather than spend the money and resources to challenge the subpoena, companies can simply turn it over or permit the government to search their records. Since September 11, AOL and Earthlink, two of the largest ISPs, have readily cooperated with the investigation of the terrorist attacks.[28] Often, ISPs have their own technology to turn over communications and information about targets of investigations. If they lack the technology, law enforcement officials can install devices such as "Carnivore" to locate the information.[29] Carnivore, now renamed to the more innocuous "DCS1000," is a computer program installed by the FBI at an ISP.[30] It can monitor all ISP email traffic and search for certain keywords in the content or headers of the email messages.

These developments are troubling because private-sector companies often have weak policies governing when information may be disclosed to the government. The privacy policy for the MSN network, an affiliation of several Microsoft, Inc. websites such as Hotmail (an email service), Health, Money, Newsletters, eShop, and Calendar, states:

> MSN Web sites will disclose your personal information, without notice, only if required to do so by law or in the good faith belief that such action is necessary to: (a) conform to the edicts of the law or comply with legal process served on Microsoft or the site.[31]

Though somewhat unclear, this privacy policy appears to require a subpoena or court order for the government to obtain personal data.

Amazon.com's privacy policy reads: "We release account and other personal information when we believe release is appropriate to comply with law . . . or protect the rights, property, or safety of

Amazon.com, our users, or others."[32] It is unclear from this policy the extent to which Amazon.com, in its discretion, can provide information to law enforcement officials.

EBay, a popular online auction website, has a policy stating that

> [it] cooperates with law enforcement inquiries, as well as other third parties to enforce laws, such as: intellectual property rights, fraud and other rights. We can (and you authorize us to) disclose any information about you to law enforcement or other government officials as we, in our sole discretion, believe necessary or appropriate, in connection with an investigation of fraud, intellectual property infringements, or other activity that is illegal or may expose us or you to legal liability.[33]

This policy gives eBay almost complete discretion to provide the government with whatever information it deems appropriate.

Truste.com, a nonprofit organization providing a "trustmark" for participating websites that agree to abide by certain privacy principles, has drafted a model privacy statement that reads: "We will not sell, share, or rent [personal] information to others in ways different from what is disclosed in this statement."[34] The statement then says that information may be shared with "an outside shipping company to ship orders, and a credit card processing company to bill users for goods and services." Personal data is also shared with third parties when the user signs up for services that are provided by those third parties. This policy, however, does not contain any provision about supplying information to the government. Further, the policy does not inform people that under existing law, information must be disclosed to the government pursuant to a subpoena or court order.

The government is also increasing information flow from the private sector by encouraging it to develop new information gathering technologies. Private-sector firms stand to profit from developing such technologies. Since September 11, companies have expressed an eagerness to develop national identification systems and face-recognition technology.[35] In addition, the federal government has announced a "wish list" for new surveillance and investigation technologies.[36] Companies that invent such technologies can obtain lucrative government contracts.

The government has also funded private-sector information gathering initiatives. For instance, a company that began assembling a national database of photographs and personal information as a tool to guard against consumer fraud has received $1.5 million from the Secret Service to aid in the development of the database.[37]

In certain circumstances, where institutions do not willingly cooperate with the government, the law requires their participation. For example, the Bank Secrecy Act of 1970 forces banks to maintain records of financial transactions to facilitate law enforcement needs—in particular, investigations and prosecutions of criminal, tax, or regulatory matters.[38] All federally insured banks must keep records of each customer's financial transactions and must report to the government every financial transaction in excess of $10,000.[39] The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 requires employers to report personal information from all new employees including SSNs, addresses, and wages.[40] The Communications Assistance for Law Enforcement Act of 1994 requires telecommunications service providers to develop technology to assist government surveillance of individuals.[41] Passed in 2001, the USA-PATRIOT Act authorizes the FBI to obtain a court order to inspect or seize "books, records, papers, documents, or other items" for use in an investigation for terrorism or intelligence activities.[42] This provision contains a gag order, prohibiting anybody from disclosing that the FBI has sought or obtained anything.[43]

All of this suggests that businesses and government have become allies. When their interests diverge, the law forces cooperation. The government can increasingly amass gigantic dossiers on millions of individuals, conduct sweeping investigations, and search for vast quantities of information from a wide range of sources, without any probable cause or particularized suspicion. Information is easier to obtain, and it is becoming more centralized. The government is increasingly gaining access to the information in our digital dossiers. As Justice Douglas noted in his dissent when the Court upheld the constitutionality of the Bank Secrecy Act:

> These [bank records] are all tied to one's SSN; and now that we have the data banks, these other items will enrich that store-

> house and make it possible for a bureaucrat—by pushing one button—to get in an instant the names of the 190 million Americans who are subversives or potential and likely candidates.[44]

Thus, we are increasingly seeing collusion, partly voluntary, partly coerced, between the private sector and the government. While public attention has focused on the Total Information Awareness project, the very same goals and techniques of the program continue to be carried out less systematically by various government agencies and law enforcement officials. We are already closer to Total Information Awareness than we might think.

The Orwellian Dangers

Although there are certainly many legitimate needs for law enforcement officials to obtain personal data, there are also many dangers to unfettered government access to information. There are at least two general types of harms, some best captured by the Orwell metaphor and others that are more fittingly described with the Kafka metaphor. I turn first to the Orwellian dangers.

Creeping toward Totalitarianism. Historically, totalitarian governments have developed elaborate systems for collecting data about people's private lives.[45] Although the possibility of the rise of a totalitarian state is remote, if our society takes on certain totalitarian features, it could significantly increase the extent to which the government can exercise social control. Justice Brandeis was prescient when he observed that people "are naturally alert to repel invasion of their liberty by evil-minded rulers. The greatest dangers to liberty lurk in the insidious encroachment by men of zeal, well-meaning but without understanding."[46]

Democracy and Self-Determination. Even if government entities are not attempting to engage in social control, their activities can have collateral effects that harm democracy and self-determination. Paul Schwartz illustrates this with his theory of "constitutive privacy." According to Schwartz, privacy is essential to both individuals and

communities: "[C]onstitutive privacy seeks to create boundaries about personal information to help the individual and define terms of life within the community." As a form of regulation of information flow, privacy shapes "the extent to which certain actions or expressions of identity are encouraged or discouraged." Schwartz contends that extensive government oversight over an individual's activities can "corrupt individual decision making about the elements of one's identity."[47] Likewise, Julie Cohen argues that a "realm of autonomous, unmonitored choice . . . promotes a vital diversity of speech and behavior." The lack of privacy "threatens not only to chill the expression of eccentric individuality, but also, gradually, to dampen the force of our aspirations to it."[48]

Freedom of Association. Government information collection interferes with an individual's freedom of association. The Court has held that there is a "vital relationship between freedom to associate and privacy in one's associations."[49] In a series of cases, the Court has restricted the government's ability to compel disclosure of membership in an organization.[50] In *Baird v. State Bar,*[51] for example, the Court has declared: "[W]hen a State attempts to make inquiries about a person's beliefs or associations, its power is limited by the First Amendment. Broad and sweeping state inquiries into these protected areas . . . discourage citizens from exercising rights protected by the Constitution."[52] The government's extensive ability to glean information about one's associations from third party records without any Fourth Amendment limitations threatens the interests articulated in these cases.[53]

Anonymity. Extensive government information gathering from third party records also implicates the right to speak anonymously. In *Talley v. California,* the Court struck down a law prohibiting the distribution of anonymous handbills as a violation of the First Amendment. The Court held that "[p]ersecuted groups and sects from time to time throughout history have been able to criticize oppressive practices and laws either anonymously or not at all." Further, the Court reasoned, "identification and fear of reprisal might deter perfectly peaceful discussions of public matters of importance."[54] The Court

has reiterated its view of the importance of protecting anonymous speech in subsequent cases.[55] From third parties, especially ISPs, the government can readily obtain an anonymous or pseudonymous speaker's identity. Only computer-savvy users can speak with more secure anonymity. Although private parties attempting to identify an anonymous speaker through subpoenas have been required to satisfy heightened standards,[56] no such heightened standards have yet been applied when the government seeks to obtain the information.

Further, beyond typical anonymity is the ability to receive information anonymously. As Julie Cohen persuasively contends: "The freedom to read anonymously is just as much a part of our tradition, and the choice of reading materials just as expressive of identity, as the decision to use or withhold one's name."[57] The lack of sufficient controls on the government's obtaining the extensive records about how individuals surf the web, the books and magazines they read, and the videos or television channels they listen to can implicate this interest.

Additionally, the increasing information flow between the private sector and the government not only impacts the privacy of the target of an investigation, but can also affect the privacy of other individuals. The names, addresses, phone numbers, and a variety of data about a number of individuals can be ensnared in records pertaining to the target.

These types of harms can inhibit individuals from associating with particular people and groups and from expressing their views, especially unpopular ones. This kind of inhibition is a central goal of Orwell's Big Brother. Although it certainly does not approach the same degree of oppressiveness as Big Brother, it reduces the robustness of dissent and weakens the vitality of our communication.

The Kafkaesque Dangers

The second general type of danger promoted by government information gathering consists of the harms routinely arising in bureaucratic settings: decisions without adequate accountability, dangerous pockets of unfettered discretion, and choices based on short-term goals without consideration of the long-term consequences or the larger social effects. These bureaucratic harms have similarities to

those I discussed earlier when discussing the Kafka metaphor, although these harms take on some new dimensions with government law enforcement bureaucracy. As in Kafka's *The Trial,* dossiers circulate throughout a large government bureaucracy, and individuals are not informed how their information is used and how decisions are made based on their data. The existence of dossiers of personal information in government bureaucracies can lead to dangers such as hasty judgment in times of crisis, the disparate impact of law enforcement on particular minorities, cover-ups, petty retaliation for criticism, blackmail, framing, sweeping and disruptive investigations, racial or religious profiling, and so on.

The most frequent problem is not that law enforcement agencies will be led by corrupt and abusive leaders, although this arguably happened for nearly 50 years when J. Edgar Hoover directed the FBI. The problem is the risk that judgment will not be exercised in a careful and thoughtful manner. In other words, it stems from certain forms of government information collection shifting power toward a bureaucratic machinery that is poorly regulated and susceptible to abuse. This shift has profound social effects because it alters the balance of power between the government and the people, exposing individuals to a series of harms, increasing their vulnerability and decreasing the degree of power they exercise over their lives.

When the Fourth Amendment was ratified, organized police forces did not exist.[58] Colonial policing was the "business of amateurs."[59] Sheriffs did not have a professional staff; they relied heavily on ordinary citizens to serve as constables or watchmen, whose primary duties consisted of patrolling rather than investigating.[60] The government typically became involved in criminal investigations only after an arrest was made or a suspect was identified, and in ordinary criminal cases, police rarely conducted searches prior to arrest.[61]

Organized police forces developed during the nineteenth century, and by the middle of the twentieth century, policing reached an unprecedented level of organization and coordination.[62] At the center of the rise of modern law enforcement was the development of the FBI. When the FBI was being formed in 1908, there was significant opposition in Congress to a permanent federal police force.[63] Members of Congress expressed trepidation over the possibility that such an in-

vestigatory agency could ascertain "matters of scandal and gossip" that could wind up being used for political purposes.[64] These concerns related to the potential dangers of the agency's information gathering capabilities, and as will be discussed later, the fears eventually became realities.

Today, we live in an endless matrix of law and regulation, administered by a multitude of vast government bureaucracies. Like most everything else in modern society, law enforcement has become bureaucratized. There are large police departments armed with sophisticated technology that coordinate with each other. There are massive agencies devoted entirely to investigation and intelligence gathering. One of the distinctive facets of law enforcement bureaucracy in the United States is that low-ranking officials exercise a profound degree of discretion, and most of their discretionary decisions are undocumented.[65]

Many factors make it difficult for law enforcement officials to strike a delicate balance between order and liberty. Among them, there are tremendous pressures on law enforcement agencies to capture criminals, solve notorious crimes, keep crime under control, and prevent acts of violence and terrorism. This highly stressful environment can lead to short cuts, bad exercises of discretion, or obliviousness and insensitivity to people's freedom. One of the most crucial aspects of keeping government power under control is a healthy scrutiny. Most law enforcement officials, however, are unlikely to view themselves with distrust and skepticism. Police and prosecutors are too enveloped in the tremendous responsibilities and pressures of their jobs to remain completely unbiased.

In short, one need not fear the rise of a totalitarian state or the inhibition of democratic activities to desire strong controls on the power of the government in collecting personal information. The Kafka metaphor more aptly captures what is harmful about these types of bureaucratic realities. The harm is that our personal data is stored within a bureaucratic system, where we are vulnerable to abuses, careless errors, and thoughtless decisions.

Leaks, Lapses, and Vulnerability. As more private-sector data becomes available to the government, there could be a de facto national

database, or a large database of "suspicious" individuals.[66] Federal governmental entities have engaged in extensive data gathering campaigns on various political groups throughout the twentieth century. From 1940 through 1973, for example, the FBI and CIA conducted a secret domestic intelligence operation, reading the mail of thousands of citizens. The FBI's investigations extended to members of the women's liberation movement and prominent critics of the Vietnam War, and the FBI obtained information about personal and sexual relationships that could be used to discredit them. During the McCarthy era and again in the 1980s, the FBI sought information from libraries about the reading habits of certain individuals. Between 1967 and 1970, the U.S. Army conducted wide-ranging surveillance, amassing extensive personal information about a broad group of individuals. The impetus for the Army's surveillance was a series of riots that followed Dr. Martin Luther King, Jr.'s assassination. The information collected involved data about finances, sexual activity, and health. In 1970, Congress significantly curtailed the Army's program, and the records of personal information were eventually destroyed.[67]

The danger of these information warehousing efforts is not only that it chills speech or threatens lawful protest, but also that it makes people more vulnerable by exposing them to potential future dangers such as leaks, security lapses, and improper arrests. For example, during the late 1960s and early 1970s, the Philadelphia Police Department (PPD) compiled about 18,000 files on various dissident individuals and groups. During a national television broadcast, PPD officials disclosed the names of some of the people on whom files were kept.[68]

Automated Investigations and Profiling. Government agencies are using personal information in databases to conduct automated investigations. In 1977, in order to detect fraud, the federal government began matching its computer employee records with those of people receiving federal benefits.[69] With the use of computers to match records of different government entities, the government investigated millions of people. Some matching programs used data obtained from merchants and marketers to discover tax, welfare, and food stamp fraud as well as to identify drug couriers.[70] This sharing of records between different government agencies, ordinarily a violation of the Privacy

Act, was justified under the "routine use" exception.[71] Computer matching raised significant concerns, and in 1988, Congress finally passed a law regulating this practice.[72] The law has been strongly criticized as providing scant substantive guidance and having little practical effect.[73]

This type of automated investigation is troubling because it alters the way that government investigations typically take place. Usually, the government has some form of particularized suspicion, a factual basis to believe that a particular person may be engaged in illegal conduct. Particularized suspicion keeps the government's profound investigative powers in check, preventing widespread surveillance and snooping into the lives and affairs of all citizens. Computer matches, Priscilla Regan contends, investigate everyone, and most people who are investigated are innocent.[74]

With the new information supplied by the private sector, there is an increased potential for more automated investigations, such as searches for all people who purchase books about particular topics or those who visit certain websites, or perhaps even people whose personal interests fit a profile for those likely to engage in certain forms of criminal activity. Profiles work similarly to the way that Amazon.com predicts which products customers will want to buy. They use particular characteristics and patterns of activity to predict how people will behave in the future. Of course, profiles can be mistaken, but they are often accurate enough to tempt people to rely on them. But there are even deeper problems with profiles beyond inaccuracies. Profiles can be based on stereotypes, race, or religion. A profile is only as good as its designer. Profiles are often kept secret, enabling prejudices and faulty assumptions to exist unchecked by the public. As Oscar Gandy observes, the use of profiling to form predictive models of human behavior incorrectly assumes that "the identity of the individual can be reduced, captured, or represented by measurable characteristics." Profiling is an "inherently conservative" technology because it "tends to reproduce and reinforce assessments and decisions made in the past."[75] Spiros Simitis explains that a profiled individual is "necessarily labeled and henceforth seen as a member of a group, the peculiar features of which are assumed to constitute her personal characteristics. Whoever appears in the lists

as a 'tax-evader,' 'assistance-chiseler,' or 'porno-film viewer' must be constantly aware of being addressed as such."[76]

Profiling or automated investigations based on information gathered through digital dossiers can result in targets being inappropriately singled out for more airport searches, police investigations, or even arrest or detention. Indeed, the federal government recently announced the creation of CAPPS II, the Computer Assisted Passenger Prescreening System, which employs computer databases to profile individuals to determine their threat level when flying. Based on their profiles, airline passengers are classified as green, yellow, or red. Passengers labeled green are subject to a normal security check; those in the yellow category receive additional searching; and those branded as red are not permitted to fly.[77] The government has not released details about what information is gathered, how people are profiled, whether race or nationality is a factor, or what ability, if any, people will have to challenge their classification.

People ensnared in the system face considerable hassle and delay. For example, in 2003, a 29-year-old member of the U.S. national rowing team was stopped at the gate when flying from Newark to Seattle. Although born in the United States, the young rower had a Muslim last name, which was probably a factor that led him to be placed on the no-fly list. When officials investigated, they cleared him, but it was too late—his flight had already left. This wasn't an isolated incident; it happened to him a few months earlier as well.[78] With no way to clear his name, he remains at risk of being detained, hassled, and delayed every time he goes to an airport.

Overreacting in Times of Crisis. The government can use dossiers of personal information in mass roundups of distrusted or suspicious individuals whenever the political climate is ripe. As legal scholar Pamela Samuelson observed: "One factor that enabled the Nazis to efficiently round up, transport, and seize assets of Jews (and others they viewed as 'undesirables') was the extensive repositories of personal data available not only from the public sector but also from private sector sources."[79] In the United States, information archives greatly assisted the roundups of disfavored groups, including Japanese Americans during World War II. Following the bombing of Pearl

Harbor on December 7, 1941, the FBI detained thousands of Japanese American community leaders in internment camps. These initial roundups were facilitated by an index of potentially subversive people of Japanese descent compiled by the Justice Department beginning in the late 1930s. In 1942, in the name of national security, about 120,000 people of Japanese descent living on the West Coast were imprisoned in internment camps. The Census Bureau prepared special tabulations of Japanese Americans, which assisted in the relocation.[80]

The acquisition of personal data also facilitated the Palmer Raids (or "Red Scare") of 1919–1920. A bomb blew up at the doorstep of Attorney General A. Mitchell Palmer's home.[81] Shortly thereafter, bombs went off in eight other cities. Letter bombs were mailed to many elites, but most were halted at the post office due to inadequate postage.[82] In a climate rife with fear of "Reds," anarchists, and labor unrest, Congress tasked the Bureau of Investigation (the organization that became the FBI in 1935) with addressing these terrorist threats.[83] Under the direction of a young J. Edgar Hoover, the Bureau of Investigation developed an extensive index of hundreds of thousands of radicals.[84] This data was used to conduct a massive series of raids, in which over 10,000 individuals suspected of being Communists were rounded up, many without warrants.[85] The raids resulted in a number of deportations, many based solely on membership in certain organizations.[86] When prominent figures in the legal community such as Roscoe Pound, Felix Frankfurter, and Zechariah Chafee, Jr., criticized the raids, Hoover began assembling a dossier on each of them.[87]

Additionally, personal information gathered by the FBI enabled the extensive hunt for Communists during the late 1940s and 1950s—a period of history that has since been criticized as a severe over-reaction, resulting in the mistreatment of numerous individuals, and impeding the reform agenda begun in the New Deal.[88] According to historian Ellen Schrecker, federal agencies' "bureaucratic interests, including the desire to present themselves as protecting the community against the threat of internal subversion, inspired them to exaggerate the danger of radicalism."[89] Senator Joseph R. McCarthy, the figure who epitomized the anti-Communism of the 1950s, received substantial assistance from Hoover, who secretly released information about suspected Communists to McCarthy.[90] Further, the FBI supplied a steady

stream of names of individuals to be called before the House Un-American Activities Committee (HUAC).[91] As historian Richard Powers observes, "information derived from the [FBI's] files was clearly the lifeblood of the Washington anti-communist establishment."[92] The FBI also leaked information about suspected individuals to employers and the press.[93] Public accusations of being a Communist carried an immense stigma and often resulted in a severe public backlash.[94] Individuals exposed as Communists faced retaliation in the private sector. Numerous journalists, professors, and entertainers were fired from their jobs and blacklisted from future employment.[95]

In short, government entities have demonstrated substantial abilities to gather and store personal information. Combined with the extensive data available about individuals in third party records, this creates a recipe for similar or greater government abuses in the future.

Changing Purposes and Uses. Information obtained by the government for one purpose can readily be used for another. For example, suppose the government is investigating whether a prominent critic of the war against terrorism has in any way assisted terrorists or is engaged in terrorism. In tracking an individual's activities, the government does not discover any criminal activity with regard to terrorism, but discovers that a popular website for downloading music files has been visited and that copyright laws have been violated. Such information may ultimately be used to prosecute copyright violations as a pretext for the government's distaste for the individual's political views and beliefs. Further, dossiers maintained by law enforcement organizations can be selectively leaked to attack critics.

Indeed, it is not far-fetched for government officials to amass data for use in silencing or attacking enemies, critics, undesirables, or radicals. For example, J. Edgar Hoover accumulated an extensive collection of files with detailed information about the private lives of numerous prominent individuals, including presidents, members of Congress, Supreme Court justices, celebrities, civil rights leaders, and attorney generals.[96] Hoover's data often included sexual activities.[97] Hoover used this information to blackmail people or to destroy their reputations by leaking it. Often, however, he did not even have to

make any explicit threats. Politicians—and even presidents—feared that Hoover had damaging information about them and would avoid criticizing Hoover or attempting to remove him as FBI director. Indeed, on one of the tapes President Nixon recorded in the Oval Office, he declared that he could not fire Hoover because Hoover knew too much information about him.[98]

We live in a world of mixed and changing motives. Data that is obtained for one purpose can be used for an entirely different purpose as motives change. For example, for several years, the FBI extensively wiretapped Martin Luther King, Jr.[99] They wiretapped his home, his office, and the hotel rooms that he stayed at when traveling.[100] Based on the wiretaps, the FBI learned of his extensive partying, extramarital affairs, and other sexual activities. A high-level FBI official even anonymously sent him a tape with highlights of the FBI's recordings, along with a letter that stated:

> King, there is only one thing left for you to do. You know what it is. You have just 34 days in which to do (this exact number has been selected for a specific reason, it has definite practical significant [*sic*]). You are done. There is but one way out for you. You better take it before your filthy, abnormal fraudulent self is bared to the nation.[101]

Hoover's motive is disputed. One theory is that King was wiretapped because he was friendly with a person who had previously been a member of the Communist Party.[102] Another theory is that Hoover despised King personally. Hoover's longstanding hatred of King is evidenced by his nasty public statements about King, such as calling King "the most notorious liar" in the nation.[103] This was probably due, in part, to King's criticism of the FBI for inadequately addressing the violence against blacks in the South, Hoover's overreaction to any criticism of the FBI, and the FBI's practice of consistently targeting its critics.[104] As David Garrow hypothesizes, the original reason that the FBI began collecting information about King was due to fears of Communist ties; however, this motivation changed once these fears proved unfounded and several powerful individuals at the FBI expressed distaste for King's sexual activities and moral behavior.[105]