“I saw my rap sheet, it's yellow with fifty-four entries. Typed up not so neat. You see your life with carbon copies. Kept in a file. . . . I got a taste for supermarket cakes. You won't find that on my yellow sheet.”

- Tillie Henderson alias Miss Bliss alias Puzzle alias Rosa P. alias Sweet-Cakes

I. INTRODUCTION

When people think of surveillance, they think of the government. They think of Closed-Circuit Television cameras (CCTVs). They think of wiretaps. They think of law enforcement, of the police, and of the FBI. The more technologically savvy among us might worry about biometric surveillance and radio-frequency identification (RFID) technology. We all think Orwell and 1984 and Big Brother. Surveillance comes from the French word for “watching over,” which is exactly what we think the government is doing: watching. This watching and listening is all reflexively—instinctively—a violation of our privacy, and it puts us on edge.

But the surveillance game is changing in the modern era. The main tool in the government’s arsenal isn’t watching or listening—it’s predicting. That is, the government now uses data collection and statistical analysis to predict how individuals will act in the future.

---

1 Colum McCann, Let the Great World Spin (2009).
3 Id.
But that’s not the only change in the surveillance game—the new development is who else is doing it, and the answer is everyone. Private companies surveil all of their users. Individuals surveil their friends, acquaintances, and even strangers. Alongside the creation of the modern surveillance state, we have created a private culture of surveillance. Technology has not only increased the government’s capacity for information collection and processing—it’s enhanced our ability to follow and monitor behavior as well. Online platforms like Facebook, Google, and Spokeo all allow us to track down personal information in profoundly new ways. Taken together, these three actors—government, corporations, and individuals—make up what I call our surveillance society.

It is no surprise that many scholars are worried about the loss of privacy in a society that privileges surveillance above much else.

Some may immediately object that what happens on Facebook is not surveillance, and, of course, in many ways they are right. My definition of surveillance is intentionally broad. Surveillance will be defined to include any collection, analysis, or processing of information about an individual or institution. Second, and related to the first, surveillance is carried out by a wide variety of actors (including the government, corporations, and individuals). Third, the government does not ultimately have to be behind the information gathering for it to constitute surveillance. That is, an individual can (and does) surveil other individuals, even if she never relays the fruits of that surveillance to other actors.
I broaden surveillance to include all types of information gathering and
information gatherers because it makes it clear that the gathering and gatherers exist and
surveil in a more complex information ecology than it may seem if the government is the
only player. Each actor in this ecology depends on and reacts to the other actors, often in
ways that mitigate some of the concerns and potential harms to individuals and society at
large.

At the same time, this paper’s scope is a narrower subset of surveillance that has
its own particular harms. Clearly, the surveillance society as a whole is rampant with
privacy violations. Many of these violations have possible legal solutions, including the
constitution, criminal law, and statutory regulation. The backbone of the surveillance
society, however, is data that we ourselves voluntarily contribute to the information
ecology (for example, through transactions with corporations, transactions with the
government, or peer interactions). Data is collected as we navigate the world (online and
off); that information is collated and attached to an identifiable user; and that information
is analyzed in order to better understand that users proclivities and future actions. By
focusing on these issues, I hope to avoid questions like whether the government can place
listening devices throughout a city.

I have other reasons for narrowing the inquiry. First, I believe that this type of
information-based surveillance that relies on voluntarily contributed information will be
the government’s (and corporation’s) most powerful means of learning about individuals

---

4 There is, of course, plenty of disagreement about whether information contributed to
unique sources should be designated as having been contributed to the wider
“information ecology.” This subject will be taken up most thoroughly in Part III of the
paper.

5 An important question, to be sure, but one for another paper.
and predicting their future behavior. Second, freely given, information-based surveillance will affect everyone in our society, whereas more traditional forms of surveillance (such as real-time observation) will likely affect a smaller subset of the population that is targeted for criminal investigation. For most people, the surveillance problem is not about mistaken rendition—those are the cases of extreme abuse and are thus atypical. Rather, for most people, the surveillance problem will be about their ability to control their reputations, to control who they are.

Third, I believe that this type of surveillance—surveillance based on information that we are responsible for providing about ourselves—presents certain unique challenges to privacy advocates. Those challenges include the fact that: (1) we initially give up the information willingly weakens the privacy interests, even if it does not vitiate them entirely; (2) we give the information to a not insubstantial group of “others,” even if we do not give it directly to the person or entity that harms the privacy; (3) this information is often available to the public at large, and, if it’s not, it tends to be governed by a contract; and (4) there are many tangible and substantial benefits that we gain through the free flow of information.

I start by providing an analytic framework for the paper: the privacy violations that inhere in a society fundamentally entwined in mutual surveillance. In Part II, I provide a summary of some of the values that support privacy protection (including dignity, autonomy, and self-determination), and I describe Daniel Solove’s taxonomy of privacy, which I rely on later to analyze the privacy harms in the surveillance society. In Part III, I describe the surveillance state, corporate surveillance, and peer-produced surveillance and offer a quick accounting of the privacy harms in each segment of the
surveillance society. I conclude each section by discussing some of the more popular proposals for legal solutions and regulation; here, I try to cast doubt on these solutions without discrediting them entirely. These discussions foreground my description of the surveillance ecology in Part IV, where I attempt to show how the ecology balances many of the important interests at stake. Although privacy advocates will certainly remain unassuaged, I tentatively conclude that the surveillance ecology affords individuals unique opportunities to shape their reputations and lead autonomous lives online.

II. PRIVACY: WHAT ARE WE PROTECTING?

Most of us hold a firm belief about the modern era: privacy is rapidly disappearing. Intuitively, many people, and especially privacy scholars, also believe that we can and should try to save privacy from its apparent fate—that if we just add legal protections, regulations, and technological enhancements we can carve out a space for privacy in the modern surveillance society. Without privacy, the argument goes, we would be continuously exposed to “the all-seeing eye of judgment, forever denied the blissful release of secret solitude in which to be reconciled with ourselves and whatever divinity we may find in the universe.”

In truth, this project began over one hundred years ago when Warren and Brandeis identified a changing privacy landscape and attempted to describe a legal regime that could protect privacy from too much encroachment. Yet, before they

---

asserted the legal protections, Warren and Brandeis had to define the right. They settled on
the now famous “right of the individual to be let alone,” but what is contained in the
modern concept of privacy is far from clear. Still, before we commit to saving privacy,
we should explore privacy’s many meanings so we have a sense of what individual and
social values are at stake.

A. Privacy’s Many Meanings

Elaborating a theory of privacy has proven notoriously difficult. As Daniel Solove
puts it, “[p]rivacy . . . is a concept in disarray. . . . [It] encompass[es] (among other
things) freedom of thought, control over one’s body, solitude in one’s home, control over
personal information, freedom from surveillance, protection of one’s reputation, and
protection from searches and interrogations.” In his inquiry, Solove ultimately identifies
six general types of privacy: (1) Warren’s and Brandeis’ right to be let alone; (2) “the
ability to shield oneself from unwanted access by others”; (3) secrecy; (4) control
over personal information; (5) “the protection of one’s personality, individuality, and
dignity”; and (6) intimacy. After identifying these six theories of privacy, however,
Solove shows why each concept on its own is incomplete and insufficient to justify
privacy protections in all its contexts.

Clearly, then, privacy serves a wide range of interests, and by invoking privacy in
such a wide variety of contexts and to vindicate such a diverse set of rights, we make it

---

8 Id. at 205
10 Id. at 13.
11 Id.
12 The list and brief descriptions of each category can be found in id. at 12-13.
hard to identify an overarching theme that unifies them all. In the United States, it is commonly understood that privacy protects both an individual’s personal information and her right to an abortion, but the reasoning behind these protections must be distinct. The former centers on one’s self-definition through a right to selectively release personal information, whereas the latter protects the individual’s right to make certain decisions (or, better, to prevent the government from making those decisions for her). Given this breadth, it is unsurprising that an overarching and coherent theory of privacy is hard to come by.

Another problem is that privacy is highly contextual. As James Whitman has argued, a society’s concept of privacy may depend more on that society’s history than anything else, and when societies derive privacy from different sources, they will come up with different legal protections. What is more, privacy is highly contextual even within a given society—certain actions may be viewed as a violation of privacy in one context, but not in another, and it may depend also on the individual’s own appetite for publicity. Because of this, societies may struggle with how it should gauge whether

---

14 CITE.
15 For a discussion about privacy as a right against government encroachment, see Jed Rubenfeld, *The Right of Privacy*, 102 HARV. L. REV. 737 (1989); J. Braxton Craven, Jr., *Personhood: The Right to Be Left Alone*, 1976 DUKE L.J. 699; and Louis Henkin, *Privacy and Autonomy*, 74 COLUM. L. REV. 1410 (1974). Although this paper focuses on privacy and its relation to information about individuals, there are some common themes to both types of privacy that infuse the rest of the discussion.
16 See James Q. Whitman, *The Two Cultures of Western Privacy: Dignity Versus Liberty*, 89 GEO. L.J. 2087 (2004) (describing how the European concept of privacy as derived from dignity and the American concept of privacy as liberty derived from their historical contexts and showing how the European and American legal systems protect different interests). Compare
17 DANIEL J. SOLOVE, *UNDERSTANDING PRIVACY* 41 (2008). (“Privacy is a product of norms, activities, and legal protections, and as a result, it is culturally and historically contingent.”).
there even is a privacy interest in the first place. As Solove puts it, “Easy distinctions such as intimate versus nonintimate and secret versus nonsecret fail to account for the complex nature of what is considered private. Privacy is a dimension of social practices, activities, customs, and norms that are shaped by history and culture.”

U.S. courts have developed several doctrines that are often either too rigid or too dependent to be of much use in determining when the law should protect informational privacy. For example, courts often draw a rigid line between what is a secret and what has been limitedly exposed to even a small group of others: if the secret has been exposed at all, it is no longer private and should not be protected by law. But in reality, “[w]e have very few secrets that are secret from everybody.” In other contexts, courts ask whether the individual had a reasonable expectation of privacy, but “[w]ithout a normative component to establish what society should recognize as private, the reasonable-expectations approach provides only a status report on existing privacy norms rather than guides us toward shaping privacy law and policy in the future.”

Despite all of these issues—despite privacy’s breadth and context-dependence—privacy advocates still tend to circle around a set of values that privacy seeks to enhance or support. At this higher level, privacy is about authorship and ownership of the self—it is about the right of each individual to create him or herself. As such, it is tied to

---

21 See, e.g., CITE CASE.
individual autonomy and to the dignity of the self. It may also be tied to liberty from
government encroachment; in that sense, it is anti-totalitarian. At its most beautiful,
“privacy is the measure of the extent an individual is afforded the social and legal space
to develop the emotional, cognitive, spiritual, and moral powers of an autonomous
agent.”23 It may be derived from moral philosophy; it may find its source in “[t]he
Kantian idea of a moral autonomy as ‘self-determination’” through which an individual
can lead a life that is “self-determined, self-authored or self-created.”24

Privacy is practically important to this development as well. It allows the
individual to choose what to reveal to whom and when to reveal it. It allows her to
portray different selves to different spheres of relations25; it allows her to maintain certain
secrets, even if revealing these secrets would disclose true information.26 But privacy
isn’t just about the individual: it is also a common good.27 It helps us shape the kind of
society we want, promoting certain rules and norms for interaction throughout our social
structure.28

These values are all important to hold in the background as we proceed through
the surveillance society. However, I believe that Solove is right that no one description of

23 FERDINAND DAVID SCHOEMAN, PRIVACY AND SOCIAL FREEDOM 13 (1992)
24 Paul Roberts, Privacy, Autonomy and Criminal Justice Rights: Philosophical
Preliminaries, in PERSONAL AUTONOMY, THE PRIVATE SPHERE AND THE CRIMINAL LAW
59 (Peter Alldridge & Chrisje Brants, eds. 2001).
25 FERDINAND DAVID SCHOEMAN, PRIVACY AND SOCIAL FREEDOM 157 (1992) (“Privacy
helps maintain both the integrity of intimate spheres as against more public spheres and
the integrity of various public spheres in relation to one another.”).
26 CITE
27 See VISIONS OF PRIVACY: POLICY CHOICES FOR THE DIGITAL AGE 5 (Colin J. Bennett &
Rebecca Grant, eds., 1999) (“Along with Priscilla Regan, we would contend that privacy
serves ‘not just individual interests but also common, public, and collective purposes.’”).
privacy can adequately capture the breadth and depth of privacy protection that we have in our legal systems. In order to refine our inquiry, then, we need language by which to analyze specific privacy violations in the surveillance society. For this language, we turn to Solove’s own taxonomy of privacy.

B. Solove’s Taxonomy of Privacy

Confronted with the failure of the theories behind privacy to adequately explain the concept in all its contexts, Solove searches for a different methodology to explain privacy. He concludes that “privacy is not defined by looking for a common denominator in all things we view under the rubric of privacy.” Rather, Solove looks to examples of actual privacy violations: “Instead of pondering the nature of privacy in the abstract, we should begin with concrete problems and then use theory as a way to better understand and resolve these problems.”

By contextualizing privacy, Solove believes that we can better understand what is at stake. He claims that “[w]e should understand the value of privacy in terms of its practical consequences. Privacy should be weighed against contrasting values, and it should win when it produces the best outcome for society.” In order to aid in the investigation, Solove developed a taxonomy of privacy—a set of fourteen generic actions

29 For example, Professor Rubenfeld has convincingly shown why privacy as protecting personhood cannot theoretically ground the U.S. constitutional doctrine surrounding decisional privacy. See Jed Rubenfeld, The Right of Privacy, 102 HARV. L. REV. 737 (1989).
30 Id. at 40.
31 Id. at 41.
32 Id. at 87.
that may impinge on individuals’ privacy.\textsuperscript{33} Seven of these privacy violations will help us better understand what interests are at stake in the surveillance society.\textsuperscript{34} They are:

- **Surveillance**: Cast as the government watching, listening to, or recording of an individual, surveillance can lead to “[t]oo much social control,” a loss of “freedom, creativity, and self-development,” and may “have a chilling effect on behavior.”\textsuperscript{35} Furthermore, “awareness of the possibility of surveillance can be just as inhibitory as actual surveillance.”\textsuperscript{36}

- **Aggregation**: Aggregation is the combination and analysis of many bits of seemingly innocuous information, which, when combined, can lead to a fuller portrait of an individual. Aggregation allows for a deeper understanding of individuals and trends, but it also “unsettles expectations.”\textsuperscript{37} Furthermore, Solove claims the picture it paints is often “incomplete” because “the data is often reductive and disconnected from the original context in which it was gathered.”\textsuperscript{38}

- **Identification**: Identification connects specific data to a specific human being. As such, “identification attaches informational baggage to people” and may “inhibit one’s ability to be anonymous or pseudonymous.”\textsuperscript{39} Identification and aggregation are related, since aggregation can lead to identification as the digital dossier gets more and more specific.

\textsuperscript{33} See id. at __.
\textsuperscript{34} The other categories, like appropriation of an individual’s image or exposure of one’s body, lie outside the scope of this paper.
\textsuperscript{35} Id. at 108-109.
\textsuperscript{36} Id. at 109.
\textsuperscript{37} Id. at 119.
\textsuperscript{38} Id. at 119-120.
\textsuperscript{39} Id. at 124-125.
• **Secondary Use**: Solove identifies secondary use as “the use of data for purposes unrelated to the purposes for which the data was initially collected without the data subject’s consent.”\(^{40}\) He claims that “[t]he potential for secondary use generates fear and uncertainty over how one’s information will be used in the future, creating a sense of powerlessness and vulnerability.”\(^{41}\)

• **Exclusion**: By exclusion, Solove means “the failure to provide individuals with notice and input about their records.”\(^{42}\)

• **Disclosure**: Disclosure describes a reputational harm “when certain true information about a person is revealed to others.”\(^{43}\) Disclosure “because it makes a person a prisoner of his recorded past. . . . [I]t can inhibit their ability to reform their behavior, to have a second chance, or to alter their life’s direction.”\(^{44}\) Significantly, “[t]he law often protects against disclosure when the information is kept secret but not when others know about it.”\(^{45}\)

• **Increased Accessibility**: Increased accessibility is just as it sounds—Solove believes there may be a privacy violation if information is more easily available through technology. Although the mere increase in accessibility can

---

\(^{40}\) *Id.* at 131.

\(^{41}\) *Id.* at 132.

\(^{42}\) *Id.* at 134.

\(^{43}\) *Id.* at 142.

\(^{44}\) *Id.* at 145. Solove also observes that “that if every person’s private life were exposed, society might change its attitudes and beliefs.” *Id.* at 144. He notes that currently “the revelation of people’s personal secrets often does damage to the individuals without having much effect on changing norms,” but it is possible that increased disclosure would alter the parameters by which we judge people’s previously private lives.

\(^{45}\) *Id.* at 145.
make people feel less secure and fear uncertainty, it is mostly an instrumental violation, since it makes other violations more likely.\textsuperscript{46}

In proceeding, this paper will use Solove’s taxonomy of privacy violations to highlight certain moments in the surveillance society that may constitute privacy violations. At the same time, however, we will keep in mind the wide variety of ultimate purposes of privacy protection identified in Part II.A above. As will be shown, the surveillance state will encroach upon individual’s private lives in a number of specific ways, and, when viewed in isolation, these encroachments will seem particularly troubling and sometimes even insurmountable. In many instances, it would be unwise or impractical for the law to address the problem head on. However, Part IV of this paper will use both the taxonomy and underlying values of privacy protection to put these violations in the broader context of the entire surveillance ecology. When viewed at this higher level, I argue that we should be deeply ambivalent about these privacy violations and potential legal solutions. Further, some of our privacy concerns will be mitigated by other aspects of the surveillance ecology, and new legal priorities will be identified. But first we need a description of the surveillance society.

III. THE SURVEILLANCE SOCIETY

A. The Surveillance State

A government engages in surveillance for many reasons. It may surveil the targets of criminal investigation. It may use surveillance (or the threat thereof) to control its

\textsuperscript{46} Id. at 150.
population. But in the modern era, more than anything else, governments surveil in order to predict the future.

With this change in scope has come a change in methodology—the bulk of a modern government’s surveillance is actually much less invasive than typically conceived. The key to mass government surveillance in the modern era is information: the government is in the business of collecting and analyzing data. By analyzing the data, the government is able to predict future action within a set of probabilities, and the state can then choose to intervene as it sees fit. In other words, the government has shifted its model of surveillance from real-time viewing to back-room computer analysis.

In the section that follows, I describe what is referred to as the surveillance state. The concept is nothing new, and I will draw extensively upon the writing of Jack Balkin and Daniel Solove, who have described different aspects of the surveillance state. I will then show how traditional modes of surveillance operate within the larger structure of the surveillance state. Once I have described the modern surveillance state, I will highlight some of the attendant privacy harms. Finally, I will discuss some potential legal solutions and attempt to show why regulating to protect privacy itself may be difficult in this surveillance landscape.

i. The Modern Surveillance State

---

47 Here, as throughout, I am concerned with the widest possible effects of surveillance on the general population. Individual instances of real-time surveillance that are used in order to catch a criminal or build a criminal case clearly relies on the actual surveillance of those involved. What I am interested in, however, is how the surveillance apparatus affects the average person.
Any surveillance story that focuses on “being watched” or “listened in on” obscures the most important part of the surveillance state: data collection and analysis. Jack Balkin has traced the origins identified the purpose of the modern surveillance state. He describes it as “a special case of the Information State,” one “that tries to identify and solve problems of governance through the collection, collation, analysis, and production of information.”\textsuperscript{48} Now “a permanent feature of governance,”\textsuperscript{49} the surveillance state was born of the welfare state and national security state. On the one hand, it is a product of the collection of tons of personal information by federal agencies that was used to deliver social services and “created a huge demand for data processing technologies to identify individuals.”\textsuperscript{50} On the other, it is the product of a post-9/11 world that is obsessed with preventing the next terrorist attack. “The National Security State . . . funded the development of increasingly powerful technologies for surveillance, data collection, and data mining, not to mention increasingly powerful computer and telecommunications technologies.”\textsuperscript{51}

In analyzing its data, the surveillance state relies on statistics and probabilities, and the power of data aggregation and data mining are profound. Modern technology permits new mechanisms of aggregation, which in turn allows the government to assemble a tremendous number of isolated data points pertaining to a given an individual. As Solove notes, “[v]iewed in isolation, each piece of our day-to-day information is not all that telling; viewed in combination, it begins to paint a portrait about our

\textsuperscript{48} Balkin, \textit{supra} note \_, at 3.
\textsuperscript{49} \textit{Id.} at 4.
\textsuperscript{50} \textit{Id.} at 6.
\textsuperscript{51} \textit{Id.}
personalities.”52 Over time, our “digital dossiers are increasingly 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.”53 These digital biographies can then be compared with other profiles, and, through statistical analysis, probabilistic predictions about the individual’s future actions can be made.

It’s this second step that makes insight into the future possibility—it’s when this information is combined and compared to other information that the government can draw conclusions. Data mining is the “application of database technology and techniques (such as statistical analysis and modeling) to uncover hidden patterns and subtle relationships in data, and to infer rules that allow for the prediction of future results.”54 Data mining “is a problem-solving tool” that “extracts predictive information from . . . databases, finding hidden patterns that may lie outside viewer expectations or be invisible on a case-by-case basis.”55 Governments use data mining to elicit these over-arching patterns and identify threats to national security.

In this system, data is king—the more the better. As Professor Balkin notes, the welfare state collects a fair amount of information itself: we frequently give up personal information in order to receive benefits back, and all those bits of information are stored

55 Furrow, supra note __, at 816.
in agencies across the federal government. This information is increasingly hosted online and made available to the general public in the form of digital public records. Technology, however, also allows the government to capture information that simply wasn’t available before. It is to these new systems of collection that we now turn.

ii. Modern Surveillance Technology

Governments have traditionally had limited resources and could not surveil the entire population all the time. This meant that they relied on the threat of watching to help control the population. Jeremy Bentham had exactly this insight when he proposed the Panopticon: the structure—where everyone can be observed at once but no one can observe the observer—is just as effective when there is no one physically present to do the observing.

One of the challenges of the modern era, however, is that new technology seems to enhance significantly the government’s ability to monitor its population in real time. The United Kingdom, for example, leads the charge on CCTVs, with around 4.2 million cameras spread across Great Britain. These cameras record in real time the actions of any who pass in front of their viewing window, and individual locales are experimenting with technology that threatens immediate repercussions. Some are working towards “an ‘intelligent’ camera that . . . is capable of picking out suspicious forms of behavior—for

---

56 See Daniel J. Solove, Symposium, Access and Aggregation: Public Records, Privacy and the Constitution, 86 Minn. L. Rev. 1137 (2002) (describing the information maintained in public records and how technology has increased how easily that information can be accessed).
58 CITE BENTHAM, FOUCALUT.
example, somebody who lingers a little longer than normal outside the window of a jewelers, or who move in a suspicious fashion.” 60 Others are focused on facial recognition technology, which has been in the works for years, even if it remains largely ineffective. 61

This sounds like an Orwellian future where we are all watched all the time. If we’re being optimistic, it sounds like a Behtamian Panopticon, “designed to give you the impression that everything you do . . . is under gaze of some watchful law-enforcement officer—when the reality is that most of the screens spend most of their time unwatched.” 62 Both accounts, however, fail to grasp the most important aspect of the new technology. These cameras are now information providers: they can be analyzed retroactively and used to predict future action. The shift, again, is from real-time observation to retroactive analysis, even when it seems like the government wants to catch you in the act.

The wiretapping story is perhaps a bit more mixed, but again the main thrust is towards information collection rather than real time surveillance. The U.S. has led the world in its efforts to listen in on conversations every day. The U.S. government started tapping communications as early as 1861, when President Lincoln approved the interception of telegraphs during the Civil War. 63 Since the 1970s, the U.S. has spearheaded the interception of international communication. 64 In the wake of 9/11 and

60 Id at 21.
61 Id. at 23.
62 Id. at 4.
63 Id. at 28.
64 1 TRANSBORDER DATA FLOWS: CONCERNS IN PRIVACY PROTECTION AND FREE FLOW OF INFORMATION 1 (Rein Turn, ed., 1979)
the global War on Terror, the U.S. has significantly ratcheted up its real-time monitoring of electronic and voice communications. Perhaps most infamously, FISA Amendments Act of 2008 permitted the government to systematically intercept electronic communications in the United States, “regardless of whether they were international or domestic communication.”65 Title II of the Act guarantees immunity for telecommunication companies who provide “assistance” to the government in this type of surveillance.66 In at least one case, this assistance included allowing the NSA to route Internet traffic through a separate room where it could be copied for analysis.67 This data was then presumably shipped back to the National Security Administration for analysis and profiling.68 Even here, then, the truth is that no one is “watching”; rather, computers are needed to conduct automated searches for particular language and information. It’s no easy task, and the vast majority of information goes unsearched. The NSA has indicated that “[e]very six hours, the agency collects as much data as is stored in the entire Library of Congress.”69 Because of the sheer amount of data coming through the agency, it is

66 H.R. 6304 §201.
67 See Jewel v. NSA; Hepting v. AT&T.
68 CITE. In some cases, the government may have used government contracts as an incentive for participation. See COHEN, supra note __, at 10 (describing how the CEO of Qwest Communications said that the government had threatened to revoke certain contracts when Qwest refused to participate in a surveillance program).
highly unlikely that most of the information is analyzed even by computers, let alone individuals.

None of this is to say that modern technology hasn’t enhanced the government’s ability to engage in actual, real-time surveillance of individuals. Indeed, just as corporate involvement was essential to the success of the FISA operations, it permeated other arrangements that involved the actual act of eavesdropping. These programs “enabled government eavesdroppers to monitor the content of millions of international telephone calls and electronic correspondence, including those where one of the parties was on U.S. soil.” And there are other forms of real-time surveillance. The U.S. government, for example, has spent significant energy mobilizing the general public. According to Jon Michaels, the state has deputized private citizens by asking them to report what they see and hear. The point, however, is that this type of surveillance is no longer the mainstay of the government’s surveillance.

Of course, we shouldn’t ignore some of the possible consequences of advances in video and audio technology simply because these classic surveillance are now geared towards information collection. But what has really changed with technology is the ability to turn this real-time surveillance where no one is actually watching into data that can be analyzed after the fact—and that can be done on everyone, especially in the context of electronic communications. Because government is able to record and preserve so much more information, they can use computers to analyze the surveillance as data

---

71 Michaels, supra note ___, at 1448.
72 Michaels, supra note __, at 1435-36.
points, rather than as individual incidents of interests. In this sense, then, old-form surveillance becomes just one of many inputs into the modern surveillance state. When combined with many other forms of data and surveillance, it allows the government to create a profile of the individual that can then be used to predict her future course of action. This type of surveillance—via data aggregation—is the present and future.

Needless to say, the surveillance state violates any number of Solove’s identified privacy interests. It engages in real-time surveillance, even if that surveillance is done primarily for other purposes, and as a means of social control, it is still effective. As has been argued, the government’s main tool is aggregation, and identification is assumed. Data mining, however, is responsible for still other violations. According to Solove, a major threat to ordinary individuals in the surveillance state is a digitized bureaucracy. As the state uses the fruits of its surveillance to make increasingly automated decisions, individuals may experience a Kafkaesque “suffocating powerlessness and vulnerability created by the . . . system’s use of personal data and its exclusion of the protagonists from having any knowledge or participation in the process.”73 Similarly, data mining excludes individuals from what the government knows about them.74

Secondary use violations are also apparent: the government uses data indiscriminately, whether it was intended to be used by federal agencies or not.75 Furthermore, despite the seemingly robust protections of the Privacy Act,76 the government regularly excludes individuals from knowing certain information about

73 SOLOVE, supra note ___, at 194.
74 SOLOVE, supra note ___, at 194.
75 More on this in Part II.B infra and corporate surveillance.
76 5 U.S.C. §552a (providing that each individual has a right to request all records pertaining to her in an agency’s possession).
themselves, while at the same time agencies significantly increase accessibility to this information by putting public records online.

Given these constant privacy violations and their associated harms, we next see if privacy can be directly protected through regulation of the surveillance state.

iii. Legal Solutions to the Surveillance State

There are a number of reasons that the law is an unlikely, inadequate, or inappropriate source of control over the surveillance state vis-à-vis harms to our privacy. Some would instinctively turn to the Constitution for protection, and the Fourth Amendment in particular. But the Fourth Amendment has been so severely limited that it certainly doesn’t apply to the most basic privacy harm we have identified—the collection and analysis of publicly available data.77 Unfortunately, a comprehensive statutory scheme doesn’t seem any more likely. First and foremost, there are practical reasons that stand in the way: lawmakers comprise the “state” in surveillance state, and the political climate makes it increasingly unlikely that they will choose to limit their own power. This is in part because it is seriously difficult to strike the appropriate balance between privacy protections and threats to national security.78 The surveillance state isn’t merely an inevitability; it will also inevitably and systematically encroach on our privacy.

Yet, even if we could pass some of the more popular statutory protections, there are good reasons to be hesitant. Take, for example, Professor Balkin’s call for institutional amnesia: he would like to see “new superstatutes . . . [that would require]

77 Balkin, supra note ___, at 20.
78 For a description on how the surveillance state is a product of the War on Terror, see COHEN, supra note ___.

22
that some kinds of data be regularly destroyed after a certain amount of time unless there were good reasons for retaining the data.”\textsuperscript{79} Although such a statute may seem uncontroversial, it creates its own aggregation problem. The conclusions we can draw from analyzing aggregated data improve the more informational inputs we have to analyze. Thus, by deleting data regularly, we run the risk of painting a less accurate picture of the individual. What is more, because the government uses aggregated information to make important decisions, we increase the likelihood that those decisions are wrong—both in the context of ensuring that the right person receives a welfare benefit or that the right person is put under criminal investigation. In other words, by requiring the regular deletion of data, we would likely exacerbate some of harms of aggregation identified by Solove.

The same problems are true of regulating or limiting information collection—the system works best when it has the largest amount of accurate information. Now, of course, one can make a strong argument that limitations on data collection and data retention are more than justified by the threat of a totalitarian state. The point here is not to definitively say that such regulation would be wrong-headed but to problematize popular legislative solutions. When combined with the practical hurdles described above, it seems ill-advised to put all of our eggs in these baskets.

As Balkin himself notes, however, there is non-privacy-based regulation that seems warranted.\textsuperscript{80} For example, we can and should ensure as transparent a system as possible that has the appropriate oversight. Solove believes that providing access to

\textsuperscript{79} Balkin, \textit{supra} note \textit{___}, at 21.

\textsuperscript{80} \textit{Id.}
information will provide little comfort for individuals because “[k]nowledge is information that has been sifted, sorted, and analyzed. The mere possession of information does not give one power; it is the ability to process that information and the capabilities to use the data that matters.” But, as will be shown in Parts II.B and II.C, private entities and individuals increasingly have to the tools that make this analysis possible. Providing the informational inputs allows individuals to challenge the government’s conclusions in meaningful ways.

In the same vein, Balkin’s idea for “create a code of proper conduct for private companies that collect, analyze, and sell personal information” also seems warranted—provided that the code of conduct focuses on letting individuals know what the government is asking for and why. We should also prioritize accuracy of the information, giving individuals the ability to readily and effectively correct inaccuracies in what is collected. Thus, we should seek to minimize Solove’s exclusion by ensuring that individuals have access to information about them.

Finally, regulators can focus on the use of the data rather than its collection and retention. For example, we should ensure that the government does not use improper categories, such as race or religion, to target individuals for further surveillance.

Despite these regulatory suggestions, privacy advocates will understandably refuse to drop their call to arms. In no way have we protected privacy itself, they don’t limit what the state has access to, whether it can aggregate, or data mine. Instead, we’ve

---

82 Balkin, *supra* note ___, at 21.
limited some uses of that information that society views inappropriate under any circumstances, and we’ve given the individual a few tools that may aid in her authoring her own identity. Although these statutory enhancements may severely limit the specter of Big Brother, they can’t completely rid us of fears about broader encroachments on autonomy, political speech, and self-determination.83

****

A huge advantage to the modern surveillance state is that it does not favor information collected by the government over information collected by other actors; the government is more than happy to appropriate information collected by corporations, who, in the natural course of business, create and maintain massive databases that track their customers’ use of their services every day. Because of this aspect of the modern surveillance state, “the line between public and private modes of surveillance and security has blurred if not vanished.”84

B. Corporate Surveillance

i. Corporate Surveillance on the Internet

Many corporations have access to individuals’ personal information, but none more than Internet companies. By tracking an individual’s web usage, one can learn what she reads before she goes to bed, what political party she identifies with, and where she likely went to dinner last week. If a corporation has access to one of the user’s social

83 For more, see Daniel J. Solove, Digital Dossiers and the Dissipation of Fourth Amendment Privacy, 75 S. CAL. L. REV. 1083, 1102 (2002).
84 Id. at 7. Corporate surveillance will be discussed in Part II.B infra.
network accounts, it may also know who is in her friend network and what are her demographics and lifestyle choices. Simply put, the Internet is a gold mine of data.

Websites interact with individuals in two fundamentally different ways: directly and indirectly. At the level that users see, Internet companies like Google, Facebook, Amazon and Twitter interact directly with the user. They provide services such as shopping, social networking, email, and video hosting, and users voluntarily upload information directly to them. Often, one company has access to user information in a number of related services. As a result, these companies have access to (and store) an incredible amount of information about their users.

Google provides an informative example. Google has developed or acquired a number of popular online services, including Gmail, Google Reader, Blogger, and Picasa.\(^85\) All of these services ask that the user to identify herself via her Google username (frequently a Gmail account), and identification offers many benefits (such as the portability of data and compatibility between the platforms). Because users tend to self-identify in order to obtain these benefits, Google is able to connect all of their activities across the services; by cross-referencing the data, the company can paint an extremely detailed portrait of the individual.

Google believes that the information it has collected is a service to the user and gives each user access to her Google information through the Google Dashboard feature. Dashboard is a detailed explanation of the information that a given user has amassed.

---

\(^85\) Gmail is Google’s web-based email platform, and Google Reader is an aggregated RSS feed that regularly collects content from sites across the Internet. Google also purchased the popular blog-hosting platform Blogger and Picasa, which allows users to upload and share photos.
across all of the Google services she uses. When I log in to Dashboard, I see my email account and nickname. I see a Blogger identity from a long-defunct blog. It shows me all of my various Google calendars, all of my contacts, all of my Google documents. But these are all surface level details; it gets more interesting when Google starts analyzing my activity. It gives me information on my Gmail usage—who I email most often and with whom I chat regularly. The coup de grace, however, is when Google compiles these usage statistics across all of its services. Its then able to paint an extremely detailed picture of an authentic social network: it tracks how I know each person (whether I email them or just follow them on Google Reader) and shows my second order connections—connections through connections. Fortunately, Google takes security seriously and “may sometimes ask you to verify your password even if you are already signed in” to other services.

This data is a valuable asset, and the analysis Google performs for users on Dashboard barely scrapes the surface of what can be gleaned from the data. Many direct service companies monetize the information they collect by selling access to their users to other companies. These third-party companies use cookies to track usage their usage across multiple sites. Users are incentivized to allow their browsers to accept cookies because they improve the functionality of a site, helping it to deliver personally tailored services; at the same time, however, the cookies upload information about the user to

---

86 Dashboard, Google, https://www.google.com/dashboard/. One needs a Google account to log in and to see her usage statistics.
87 Id.
88 Cookies “are text files written to a user’s hard disk without his knowledge [that] track our clickstreams and read data stored on our computer: user ID, password, preferences, lists of sites visited and perhaps more.” Raymond Wacks, Privacy Reconceived: Protecting Personal Information in a Digital World, in LAW, INFORMATION AND INFORMATION TECHNOLOGY 88 (Eli Lederman & Ron Shapira, 2001).
companies that then aggregate and analyze the information, which can be used to deliver targeted advertisements.\textsuperscript{89} One study focusing on social networks concluded that social network users “are vulnerable to having their [social network] identity information linked with tracking cookies.”\textsuperscript{90} This, in turn, makes it possible for third-party aggregators to associate the social network identities with past and future actions on the Internet—and the user is identified whether or not the data is technically anonymized.\textsuperscript{91} Another informal study found that the site CareerBuilder.com allowed ten different tracking companies to place cookies on the user’s browser.\textsuperscript{92} Smartphones are even worse: the Wall Street Journal found that over half of the applications tested transmitted the phone’s unique identification number, and the popular application Pandora transmitted the age, gender, location and ID number of each phone owner to multiple advertising companies.\textsuperscript{93}

Yet, as nominally creepy as this may seem, none of it is necessarily illegal: we voluntarily share this information with these companies, and that relationship is documented via contractual agreement.\textsuperscript{94} Some companies may be violating their own privacy policies, but most are amassing, cataloguing, and maintaining this information for

\begin{flushleft}
\textsuperscript{89} Id.


\textsuperscript{91} Id.


\textsuperscript{94} CITE
\end{flushleft}
legitimate business purposes. In fact, the ability of these companies to monetize user data is what drives their profitability and sustains most of the innovation on the Internet.

Of course, having all of this personal data out there can be a bit daunting for users—especially since court doctrine by and large would appear to allow these companies to sell and to trade this information at will.\(^95\) Of course, this is how the government reenters the picture: what makes the government’s new system of surveillance so effective is that it doesn’t care about the source of the information, as long as it is reliable. Therefore, as long as information is being collected and retained by private companies, the government is happy, since it can call on the private sector when it needs further informational inputs.

That the data eventually makes its way into government hands is by no means guaranteed; in fact, the vast majority of information likely never makes it into government databases. Different companies will also react differently to government requests for information based on what services they offer. Direct service companies, like Google, Twitter, and Facebook, all maintain a law enforcement exemption: they reserve the right to give up a user’s information if “it is reasonably appropriate to comply with a law, regulation or legal request” or if the company has a “good faith” reason to believe that compliance is required by law.\(^96\) Facebook even has an automated form for streamlining legal requests.\(^97\) It is this aspect of these major cites that have led Wikileaks’

\(^{95}\) CITE.
\(^{97}\) Facebook, Legal Inquiries, https://www.facebook.com/help/contact.php?show_form=legal_inquiries
Julian Assange to say that “Facebook . . . is the most appalling spying machine that has ever been invented. Here we have the worlds most comprehensive database about people, their relationships, their names, their addresses, their location their communication with each other, their relatives . . . all accessible to US intelligence.”

Yet, these companies actually have strong incentives to maintain their users’ privacy: they need you to keep contributing information to their sites so they can continue to profit off of advertising. They are probably unlikely to give up information voluntarily if doing so would negatively impact users and possibly chill their participation on the sites. It is perhaps unsurprising, then, that Google and Twitter have both fought government subpoenas for personally identifying information,\(^99\) and Facebook’s general counsel has insisted that he’s “itching for a fight.”\(^100\)

As the information moves progressively further from those companies that interact directly with user, there is less accountability. Since these companies never come

\(^98\) Julian Assange on Russia Today, May 2, 2011, http://rt.com/news/wikileaks-revelations-assange-interview/. Others have gone further, documenting an actual link between companies—including Facebook—and politically powerful people and corporations. See, e.g., COHEN, supra note __, at 53-54 (describing Facebook’s connection to Accel Partners, a group that also helped create the Defense Advanced Research Projects Agency). It is unlikely, of course, that Facebook is feeding information directly to the government.

\(^99\) See Cohen, supra note __, at 51 (noting that Google refused to comply with a subpoena that ordered Google to turn over general search information when the government was investigating child pornography on the Internet); and Ryan Singel, Twitter’s Response to WikiLeaks Subpoena Should Be the Industry Standard, WIRED, Jan. 10, 2011, http://www.wired.com/threatlevel/2011/01/twitter/ (describing Twitter’s successful challenge to a “gag order” that prevented Twitter from notifying affected users about the government’s request for account information).

in direct contact with users, they may feel free to sell the information to various parties, including the government. ChoicePoint was perhaps the most infamous: in addition to selling information to other corporations, ChoicePoint set up a specific interface just for law enforcement and tailored a service called AutoTrack to their government clients’ needs.\(^{101}\) ChoicePoint no longer exists because it was bought by Reed-Elsevier, the parent company to Lexis-Nexis. Reed-Elsevier also trades in information and “conduct[s] much business with government agencies, officials and employees as customers, authors, editors, and the like.”\(^ {102}\)

Regardless of whether the information actually makes it to the government, it is clear that corporate surveillance leads to many of Solove’s violations of privacy. Like the government, aggregation is the name of the game for the private sector—the chance to analyze the vast troves of data is what makes the data valuable in the first place. The data is clearly identifiable by the direct service companies (Google’s Dashboards shows us that). What is more, third-party companies can also often identify the user through the complex, aggregated picture they collect, even though the direct service companies often “anonymize” the data. Clearly, then, the collection of this data doesn’t just increase the chance of secondary use—that’s the very purpose, and even the government may end up having access to the data for its own surveillance purposes. In this sense, the data

---

\(^{101}\) See Chris Jay Hoofnagle, *Big Brother's Little Helpers: How Choicepoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement*, 29 N.C. J. INT’L L. & COM. REG. 595, 595-96, 611 (2004); Choicepoint, Epic.org, http://epic.org/privacy/choicepoint/ (describing the variety of information and services that ChoicePoint sold to the government, including “‘wildcard searches,’” which allows law enforcement to “obtain a comprehensive personal profile in a matter of minutes” with only a first name or partial address}).

collection also increases accessibility, since this is information that the government would not collect on its own. It also increases the risk of disclosure, since no database is a secure database. Users are also routinely excluded from their data, despite Google’s efforts to give users some access. Finally, corporations are sometimes even asked to help the government surveil individuals in real-time, as shown in Part II.A above.

Given the proliferation of privacy harms caused by corporate surveillance, many want to regulate information collection and distribution, as well as the security of databases. The next section describes some of the more popular suggestions and analyzes the merits of the proposed regulations and legal regimes.

ii. Legal Structures and the Regulation of Corporate Surveillance

There are two main strands of analysis in the literature surrounding corporate data gathering. First, scholars may focus on the collection of information, its subsequent retention, and what happens if the database in which the data is stored is breached. Second, scholars may focus on whether we should regulate corporations when they attempt to share legally obtained and truthful information about their customers.

Scholars do not always approach these issues from the perspective of consumer privacy. In fact, there is a strong conceptual argument that transactional data should not be private since “[t]hese interactions are an unavoidable incident of social life.” In some respects, “we enjoy more privacy in this regard than did any of our ancestors. Historically, most people lived in, or conducted their commercial transactions in, very

---

small communities.” At the same time, transactions have fundamentally changed: whereas a shop-keeper might know your identity and what you purchase, a website can collect much more information about your financial past and future. In other words, the individual no longer controls what information the shop-keeper has access to since most of the studying occurs without her explicit awareness. Furthermore, although the point about privacy is perhaps conceptually important, it is unlikely to affect the legal discussion—the author of the above quotes uses her observation to push for a control paradigm, which would also afford the consumer some protection.

Returning to the two strands of analysis, the first strand—about the collection, retention, and security of data—is largely outside the scope of this paper. We can and do regulate data security. At least 10 states impose obligations on companies to provide data security measures, and many more states have security breach notification laws. Scholars have proposed additional possibilities for control, including imposing a tort-like structure or analogizing to product liability. These analogies also tend to confront the issue of cybersecurity and combat identity fraud—relatively uncontroversial stances. Proposals for data deletion laws, however, still seem problematic, since they create the same problems that we saw in the context of government regulation above. Any deletion

\[104\] Id.

of information makes the surveillance system operate more inefficiently and less accurately, which can lead to a worsening of some of the harms that we are seeking to protect (like bureaucratic typecasting). These harms occur in the private sector as well, where an individual might be turned down for a financing plan because her aggregated dossier is inaccurate. Again, one could easily argue that we should trade these inaccuracies for the privacy protection afforded. My point here is not to carry the day; rather, I want to problematize the regulation so we are inclined to look elsewhere first.

The second strand of analysis—about companies that share lawfully collected information with third parties—is more complicated. First, there is a threshold constitutional question about whether regulations would unconstitutionally infringe on the corporations free speech rights.107 Many of these issues have been teed up in Sorrell v. IMS Health,108 which tests the limits of private data mining. The issue pending before the Supreme Court is “how far the government can go to restrict the commercial use of information gathered about the billions of drug prescriptions that doctors write every year.”109 Among the many issues is whether the “case [is] about access to non-public information, or . . . about censorship of a disfavored message.”110 The statute in question presents a number of interesting questions for the Court, although it may ultimately be

107 Compare Eugene Volokh, Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People from Speaking About You, 52 STAN. L. REV. 1049, 1122 (2000) (arguing that “restrictions on speech that reveals personal information are constitutional under current doctrine only if they are imposed by contract, express or implied”); and Neil M. Richards, Reconciling Data Privacy and the First Amendment, 52 UCLA L. REV. 1149, 1189-90 (2005) (disagreeing with Volokh and arguing that both data collection and use regulation would “fall outside the scope of the First Amendment.”).
108 Supreme Court Docket No. 10-779.
110 Id.
doomed by its apparent viewpoint discrimination and failure to achieve its ends—it only
prohibits the use of the data mining information in marketing contexts (despite the fact
that the data is used for other health purposes).111

Still, the First Amendment hardly settles the question here—we can and do
regulate certain types of collection and uses of personal data. Numerous federal statutes
regulate data collection and use in particular contexts.112 HIPAA, for example, contains a
privacy rule that regulates the disclosure and use of information held by covered
agencies.113 But these protections do not cover many of the privacy violations we have
described above, and would-be litigants have had to find other legal arguments. A good
example is the DoubleClick class action,114 where the plaintiffs sued under Titles I and II
of the Electronic Communications Privacy Act.115 The Title II claim was based on the
argument that the cookies that DoubleClick placed on the users browser consisted of
unauthorized access to stored electronic communications within the meaning of the state.
The court, however, held that the websites were users within the meaning of the statute,
and thus the access was in fact authorized.116 The Title I wiretap claims failed for the

111 See Andrew Tauber, After the Argument: Sorrell v. IMS Health Inc., MAYER BROWN,
(describing why Tauber believes the Court will ultimately hold the law unconstitutional).
and Video Privacy Protection Act, 18 U.S.C. §2710. For a more complete list of privacy
regulations, see MARC ROTENBERG, THE PRIVACY LAW SOURCEBOOK 1999 (1999). For a
description of some of these privacy statutes, see JACQUELINE KLOSEK, THE WAR ON
PRIVACY 1-32 (2007).
113 45 C.F.R. §§160, 164(A) and (E).
116 In re DoubleClick, Inc., 154 F. Supp. 2d at 508-09.
same reason—the websites were a party to the communication, and they consented to the tapping.\textsuperscript{117}

The DoubleClick litigation is informative for another reason, however—it eventually settled based on the contractual promises DoubleClick had made to its users. Contract is an important and viable tool for self-regulation in corporate surveillance. It is true that private individuals often ignore contractual language and “the fact that most people do not read most contracts most of the time makes contract law and imperfect tool.”\textsuperscript{118} Yet, there is reason to put our faith in contract. First, “privacy polices were found to have become more developed in their privacy promises to users across time.”\textsuperscript{119} Second, contract litigation only takes a few savvy plaintiffs to bring a case—an that case can shape industry practice. Such was the resolution to the DoubleClick litigation. There, “[t]he settlement reached in the litigations demonstrate the constant interplay between contract law and privacy issues.”\textsuperscript{120} Third, despite the increased restrictiveness of the terms of use,\textsuperscript{121} courts have developed “a sliding scale of terms-of-use enforceability” that takes into account whether the contract is a clickwrap agreement (where the user has to affirmatively indicate agreement and are often hele enforceable) or a browsewrap

\begin{flushleft}
\footnotesize
\textsuperscript{117} Id. at 514.
\textsuperscript{119} Andrea M. Matwyshyn, \textit{Mutually Assured Protection}, in \textit{SECURING PRIVACY IN THE INTERNET AGE} 74 (Anupam Chander, Lauren Gelman, & Margarit Radin, eds., 2008).
\textsuperscript{120} Sobel, Petroulakis, & Dixon-Thayer, \textit{supra} note __, at 63.
\textsuperscript{121} Matwyshyn, \textit{supra} note __, at 74.
\end{flushleft}
agreement (which requires no affirmative consent from the user and will likely not be enforced).\footnote{122 Id. at 78.} Fourth, contractual restrictions on speech pass constitutional muster.

None of this is to say that contract is the perfect solution to the corporate surveillance solution, and there are many good practices that could be developed or required in order to increase the parity of the bargaining parties. Once again, transparency is important.\footnote{123 See Sobel, Petrulakis, & Dixon-Thayer, supra note __, at 66.} Another scholar has proposed beefing up contract by awarding restitution if a site “improperly” handles or uses personal information and is unjustly enriched as a result.\footnote{124 Marcy E. Peek, Beyond Contract, in SECURING PRIVACY IN THE INTERNET AGE 140 (Anupam Chander, Lauren Gelman, & Margarit Radin, eds., 2008).} Still, contract suffers from the problems noted above. But part of the appeal of the contractual solution is its simplicity vis-à-vis other proposals. Some scholars propose an ownership model of personal data,\footnote{125 CITE.} but it would create certain ambiguities. For example, what counts as transactional data (owned by both parties) versus personal data? Second, we might be hesitant to create a new property right in true facts about a person. One could also imagine a tort analogy here, where a site might be held liable under a products liability framework for failing to deliver on consumer expectations about privacy. This argument will be dealt with more significantly in Part III.C.iv \textit{infra}; for now, it is worth it to observe quickly that it could be difficult to identify (a) consumer expectations of privacy and (b) what makes for a better design. Particularly on the second point, these sites have multiple purposes, including facilitating the transaction at hand, and it is unclear that a poor design leads to a decipherable harm.
Before moving on, it is important to see that there are a number of substantial practical reasons we may not want to limit or regulate companies that want to share legally obtained information from their users. First, data aggregation and data mining has huge benefits, especially when done by the private sector. For example, in the health sector, “data mining [may reveal] hidden sources of bad patient outcomes or substandard performance generally in the hospital setting.”\textsuperscript{126} Despite the fact that aggregation often identifies the individual, we may still opt to allow data mining for its own substantial benefits. Second, restrictions are often economically inefficient, insofar as they prevent other transactions from occurring.\textsuperscript{127} Third, even some presumptively “nefarious” uses—such as targeted marketing—have their benefits. Better marketing should create a more efficient market, delivering desired products to consumers, and we have laws ready to deal with companies that use their advertising to confuse consumers. Finally, and most importantly, the monetization of personal data is the backbone of the Internet. Much of the innovation that we know and love is sustained by the sharing and, yes, sale of personal information. We should be reticent to interfere too much in that model.

Again, none of these problems are meant to definitively discredit regulation of online data collection, retention, and trading by private companies. Rather, I hope I have problematized the case for regulation. In Part IV, we will see how our society’s system of surveillance may regulate itself enough to assuage many of the fears that have been left unresolved here. First, however, we turn to our next actors in the surveillance ecology.

\textsuperscript{126} Barry R. Furrow, Data Mining And Substandard Medical Practice: The Difference Between Privacy, Secrets and Hidden Defects, 51 VILL. L. REV. 803, 810 (2006).
\textsuperscript{127} Raymond T. Nimmer, Contracts, Markets, and Data Control, in SECURING PRIVACY IN THE INTERNET AGE 140 (Anupam Chander, Lauren Gelman, & Margarit Radin, eds., 2008).
One of the questions lurking in the background has been how corporations get all of this information on us. Much of it, as indicated, has to do with online transactions. At the same time, there are a special set of online services whose business is to collect and disseminate personal information—and we participate willingly. These services are vehicles for what I call peer-to-peer surveillance, and it is to them we now turn.

C. Peer-to-Peer Surveillance

Most would stop the analysis after a description of the surveillance state, but there’s another player in the modern surveillance game—each one of us. The rise of a culture of surveillance among individuals parallels the rise of the surveillance state. As technology has developed, particularly on the Internet, we have been offered new and powerful means of private surveillance.

Surveillance here means the same thing as it did for the surveillance state: the collection, aggregation, and analysis of information about individuals and institutions. Some may object to the use of the word “surveillance” in the private contexts that follow, but I use the term as an analytic tool: it highlights the parallels between the surveillance state and peer surveillance that I believe have been under-analyzed. There are of course many other terms that could be used to describe the technology that follows and the interactions that accompany them. Many of these terms are, on their face, quite a bit more positive in their connotations (think “social networking”). Perhaps surprisingly, my point in using the term “surveillance” is not to cast these technologies in a negative light, or even to be sensationalist; rather, it is to show that the interaction that takes place on these platforms mirrors in fundamental ways the interaction in the surveillance state. By showing the parallels, I hope to pull both peer surveillance and the surveillance state to a
more reasonable middle: peer surveillance is more problematic than it may seem, while the surveillance state is maybe less nefarious than most assume. Neither, however, should elicit doomsday scenarios for privacy hawks.

There is a second reason I use the term surveillance for both the surveillance state and these peer surveillance platforms. As will be argued in Part IV, many of the fears of privacy advocates are alleviated—as best they can be—by the interaction between the two models of surveillance.

I use three specific platforms as a window into the surveillance culture: Facebook, Spokeo, and Google (approached primarily through the Google Profile feature). Only Facebook has been well explained in pre-existing literature; the other two are less prevalent, but, I believe, essential to the future development of the web. From each of these platforms, we gain information about friends, acquaintances, and strangers. We access this information from different sources—the individual itself, her friends, her school or employer, even anonymous third parties. We use this information to learn about the individual, to paint a portrait of her over time, to make social connections, and, in many cases, to make decisions that affect her and our lives.

i. Facebook

Privacy scholars have rightfully focused on social networks, examining whether users walk the line between their social urges to participate and their desire for privacy.128 Given its prominence in the social networking world, it is no surprise that Facebook has

been the focus of at least a few of these inquiries. Facebook currently has over 500 million users, over 50 percent of which log in on any given day. Users spend 700 billion minutes per month on Facebook, or 1400 minutes per user—which means that the average user spends just under one day per month on Facebook. Access to Facebook is only getting easier: around 250 million active users access Facebook through a mobile device.

Getting onto the site is just the beginning. The average user creates 90 pieces of content each month and is connected to 80 community pages. She has 130 friends which gives the user access to roughly 130 news feeds, the associated personal information, pictures, etc. (depending, of course, on her friends’ privacy settings and participation rates). Just as important, the user can broadcast her information and updates to at least 130 friends—or far wider than that, if she wants. In total, “over 30 billion pieces of content (web links, news stories, blog posts, notes, photo albums, etc.) are shared each month.”

129 See, e.g., James Grimmelmann, Privacy as Product Safety, 19 WIDENER L.J. 793 (2010); James Grimmelmann, Saving Facebook: A Response to Professor Freiwald, 95 IOWA L. REV. BULL. 13 (2009); James Grimmelmann, Saving Facebook, 94 IOWA L. REV. 1137 (2009).

130 All statistics were drawn from Facebook itself. Press Room, https://www.facebook.com/press/info.php?statistics (last visited __). A couple of caveats about the usage statistics. First, it is unclear whether Facebook is able to distinguish between people who are logged in to the site in the background (for example, by leaving Facebook open in a tab in her browser) and those who are actively using their site. Second, there is surely no “average” Facebook user—some users spend much more time on Facebook than the average, and some much less. I use the term as a descriptive tool to show just how much time people in general spend on Facebook each month: that is, 500 million people spending almost a day each per month on the site.

131 Id.
By what is Facebook? Why do users flock to it? And why are users seemingly so willing to give up their privacy?

a. How a Desire To Be Social Creates a Tool For Social Surveillance

James Grimmelmann has gone to great efforts to explain the tension between Facebook’s social roots and privacy. His view of Facebook (and mine) is that it is a social place where people come together to create and to strengthen social connections, much in the way that they would “in real life.”

According to Grimmelmann, Facebook users crave the social definition and self-creation that the site engenders:

Facebook provides users with a forum in which they can craft social identities, forge reciprocal relationships, and accumulate social capital. These are important, even primal, human desires, whose immediacy can trigger systematic biases in the mechanisms that people use to evaluate privacy risks.

Significantly, “these social urges can't be satisfied under conditions of complete secrecy. Identity performance requires an audience; relationships are impossible without others; community is a public.” These “social dynamics . . . also cause people to misunderstand [the privacy] risks.”

At the same time, however, Facebook is a powerful vehicle for peer surveillance. Every day, it provides access to individual episodes that we can follow in (almost) real

---

133 Grimmelmann, Saving Facebook, supra note __, at 1151.
134 Id. at 1159-60.
135 Id. at 1160.
time through status updates, purchase information, “likes,” etc., all of which conveniently appear in our News Feed. Furthermore, since all of these episodes are attached to individuals, we can aggregate what we learn over time, combining it with the hordes of other personal data available on users’ profiles. And of course the information we learn isn’t limited to one user—users regularly contribute information about others (picture-tagging being perhaps the most salient). The kicker is that all of this information is surprisingly portable: not only can users copy images and text, but they can also have their News Feed ported into an RSS feed, which can then be shared easily with non-Facebook users. The only limitations on what we can see are the users’ own willingness to contribute content and specific users’ privacy settings. But most users contribute a lot of content and set minimal privacy settings, if any.

How does Facebook stock up against Solove’s taxonomy of privacy? It creates the potential for at least six of the seven violations described in Part II.B. Through the News Feed, it allows us to follow up-to-the-minute status reports on an individual’s whereabouts and activities, which is as close to a mass machine for surveillance (as Solove has defined it) as we can find online. All of the information is conveniently associated with a particular profile, with images, personal information, and friend networks available in one place. All of this information is aggregated for us over time, so we may observe the evolution of this person’s social life online. The potential for secondary uses are limitless: data may be used in marketing, by the government, or by other individuals who pass the photo or tidbit along to others. With these secondary uses

---

136 RSS stands for Really Simple Syndication; the feeds allows users to have sites forward new content directly to the user. See RSS, Wikipedia, http://en.wikipedia.org/wiki/RSS.
comes the great potential for embarrassing disclosure. Last but not least, none of this information would be available so easily if Facebook didn’t exist.

But Facebook is merely a platform, one that people have to (by and large) voluntarily use. Why, then, do users give up all this information? Grimmelmann has already provided the insight: to be social one must conduct at least some of her life before an audience. There are, however, other reasons as well. Lauren Gelman persuasively argues that it stems from the user’s desire to capture her entire social network: “Internet users are calculating that they are unlikely to identify a priori all the people they intend to reach with their posts because their social network is undefined.”137 Because of the way these sites (and the Internet at large) are set up, users are choosing to make information widely available under the (often false) assumption that “the information they post will only be accessed by people for whom it is written.”138 The problem, then, is that social networks have a “blurry edge”—and the fear of missing out on some additional social connections encourages the user to overshare.139

This urge to capture their entire, unseen social network helps explain why so few users take advantage of Facebook’s robust privacy offerings.140 After only a couple of

---

137 Gelman, supra note __, at 1317.
138 Id. at 1318.
140 Compare Grimmelmann, Saving Facebook, supra note __, at 1185-86 (discussing several studies that show how little users know about or utilize Facebook privacy protections); and id. at 1184-85 (highlighting Facebook’s “staggeringly comprehensive” set of privacy options”).
clicks, users can view a page on “Controlling How You Share.” The page contains an abundance of information on how a user can control her content and set up different spheres of friends with varying levels of access, among other mechanisms. The page wants the user to get the right message: “Facebook is about sharing. Our privacy controls give you the power to decide what and how much you share.” Facebook even goes out of its way to warn users of what they can’t control. For example, you can remove a tag so that it isn’t formally associated with your profile, but “the owner of a photo can still share that photo with people you’re not friends with.” The site also offers several video tutorials that walk users through the various controls.

In the end, we’re left where we started: “The fact is, there's a deep, probably irreconcilable tension between the desire for reliable control over one's information and the desire for unplanned social interaction.”

What should not be lost, however, is how much we love “surveilling” our friends and acquaintances. The privacy discussion (and our discussion so far) focuses on the individual user who is deciding whether or not to contribute content. But there would be no decision needed if Facebook’s other users didn’t have such an appetite for the

---

141 Controlling How You Share, Facebook, https://www.facebook.com/privacy/explanation.php (last visited __). See also Help Center, Privacy: Privacy Settings and Fundamentals, https://www.facebook.com/help/?page=839 (last visited __) (answering a number of specific privacy questions like “[h]ow can I control who can find me in Facebook search?”).

142 Id.

143 Id.

144 Ironically, when I clicked on the first video, I was told that “[t]his video either has been removed from Facebook or is not visible due to privacy settings.” The other two videos worked fine.

145 Grimmelmann, Saving Facebook, supra note __, at 1185.
information in the first place. By and large, this surveillance is harmless and social, but it also exposes the contributing users to potential, sometimes dramatic privacy harms.146

For better or worse, social network users often choose to sacrifice privacy at the altar of sociality. The site allows friends, acquaintances, and often strangers to surveil just about anyone they want. Facebook does, however, have certain built in limitations. The user can control her privacy settings and what she chooses to contribute to the site. Furthermore, she can often times exert social pressure on her friends and acquaintances about what they post about her. Facebook’s tools also allow her to know what’s out there and how much of it is directly linked to her profile. Finally, she also has some power over Facebook the company—they need her to keep coming back to the site and to contribute content so they can continue to profit.

The next peer surveillance tool, however, subverts some of this control. It combs the public web in the hopes of providing a complete picture of an individual’s public web presence, merging social spheres that some may work to keep separate.

ii. Spokeo’s Deep Web Crawling

Head to Spokeo.com, and the first thing you see is the site’s own one-line description: “Not your grandma’s white pages.”147 Undeniably true, Spokeo is one of many personal information aggregators—it combs the public web for information about

146 Grimmelmann, Saving Facebook, supra note __, at 1140 (“Over a hundred million people have uploaded personally sensitive information to Facebook, and many of them have been badly burnt as a result. Jobs have been lost, reputations smeared, embarrassing secrets broadcast to the world.”); Grimmelmann, Accidental Privacy Spills, supra note __, at 3-6 (describing an email that was forwarded to progressively wider audiences and great embarrassment to the author).
147 Spokeo, http://www.spokeo.com/ (last visited __).
an individual and aggregates it all in one place. All the user has to do is decide who to
search for and then plug in a name, email, phone number, or web username. There’s even
a section designed to help you search for your “friends” where you can upload all of your
email contacts at once.\footnote{Spokeo, http://www.spokeo.com/friends (last visited ___). It seems that Spokeo does
not use your contact information to supplement its searches of the public web; rather, it
only stores that information in your account. See Privacy Policy §2, Spokeo,
http://www.spokeo.com/blog/privacy/ (last visited ___) (describing what information
Spokeo stores and how it uses that information).} After searching for my own name, Spokeo returned for free my
address, a picture of the building via Google maps, my father’s name, and my sister’s
name (it wisely knew I had a mother, but could not identify her). Spokeo notes that the
full report may include an estimated income, photos, videos, hobbies, lifestyle, and
more.\footnote{I considered paying for the full report for the purposes of this paper but ultimately
decided that I did not want my credit card information going to Spokeo itself.}

Spokeo doesn’t even pretend to be about social networking. Go to the search by
“friends” page, and Spokeo promises that you can “[s]can your email contacts and
uncover surprising facts on everyone.”\footnote{Spokeo, http://www.spokeo.com/friends (last visited __).} To hammer home the message, the page
features a picture of a game of telephone where the last person to receive the information
is clearly surprised on learning the dirty secret. This is a place where you can learn about
gossip; this is a place where you can surveil your peers.

Perhaps surprisingly, however, Spokeo doesn’t reveal anything that isn’t already
“public” somewhere on the web. According to its privacy information page, “Spokeo
aggregates publicly available information from phone books, social networks, marketing

\footnote{\textsuperscript{148} Spokeo, http://www.spokeo.com/friends (last visited ___). It seems that Spokeo does
not use your contact information to supplement its searches of the public web; rather, it
only stores that information in your account. See Privacy Policy §2, Spokeo,
http://www.spokeo.com/blog/privacy/ (last visited ___) (describing what information
Spokeo stores and how it uses that information).
\textsuperscript{149} I considered paying for the full report for the purposes of this paper but ultimately
decided that I did not want my credit card information going to Spokeo itself.
\textsuperscript{150} Spokeo, http://www.spokeo.com/friends (last visited __).}
surveys, real estate listings, business websites, and other public sources.” Because it “does not originate data or publish user-generated content,” Spokeo is analogous to a search engine. But what separates Spokeo is how deep they are willing to look: Spokeo harnesses “deep web” technology and algorithms that allow it to scan a tremendous number of non-traditional information sources including interactive databases. For the most part, search engines have not tried to interact with databases that require the user to contribute information in order to tell the database what to look for. Spokeo and other deep-web searchers, however, have created algorithms that allow the search engine to tap into the information contained in these databases by guessing probable search terms. Spokeo’s model is probably relatively simple: it most likely “guesses” the individual’s name, email address, phone number, etc. in these databases. Other companies, including Google, are trying to use deep web technology to create a more accurate picture of all the information on the web at any given moment. Instead of merely returning the results of these deeper searches, Google “analyzes the results and develops a predictive model of what the database contains.” The possibilities are enormous if technology can “automat[e] the process of analyzing database structures and cross-referencing the results.”

---

151 Privacy, Spokeo, http://www.spokeo.com/privacy (last visited __).
152 Privacy, Spokeo, http://www.spokeo.com/privacy (last visited __).
155 Id.
Unlike Facebook, Spokeo and other personal information aggregators really seem to have one identity: they are vehicles for peer-to-peer surveillance. Although the technology is far from perfect (another aggregator still identifies me as a college newspaper reporter\textsuperscript{156}), they exist to help you learn details about people you know or know of—and Spokeo, at least, even seems to glorify in its role as facilitator. Intuitively, many experience Spokeo’s aggregation of information as a violation of privacy. If we refer back to Solove’s taxonomy of privacy, Spokeo threatens at least five of the seven violations identified in Part II.B. By aggregating information from so many different sources, it can provide a nearly complete picture of an individual in a matter of seconds, a portrait that it would take a human being many hours to cull together, if she could recreate it at all (clearly increasing accessibility to the information). Furthermore, by identifying the individual, this aggregation threatens to collapse an individual’s different reputational spheres; whereas some people may welcome Spokeo’s ability to connect her personal blog to her employment, many others would not. Because of Spokeo, however, individuals can no longer rely on others simply failing to make the connection themselves. Spokeo itself is making a secondary use of the information that users may not have originally contemplated when they put the information online, and it flaunts the fact that it may disclose embarrassing details about an individual.

Yet, despite the apparent privacy harm, it is unclear that Spokeo is doing anything wrong. Spokeo searches only the public web, and it doesn’t control the underlying information itself. Spokeo even offers to help you control your online reputation by

\textsuperscript{156} zoominfo, http://www.zoominfo.com/search (last visited __). To be fair, zoominfo is geared towards business communities, and I have yet to enter a permanent workforce.
removing your listing from its site and other personal information aggregators, but “you won’t be able to stop the flow of your information from the sources that provided Spokeo with the data in the first place.” The problem, if there is one, is how much information makes it to the public web. As mentioned above, Facebook profiles are searchable through Google—and the search result includes a small number of the user’s Facebook friends. Because the search returns different friends to different searches, “it’s possible for Google or anyone else to figure out pretty much all your friends [by repeatedly pinging public profiles].” The user has few self-help remedies as well: although a user can make her Facebook profile unsearchable with relative ease, trying to remove all of your personal data from the sites searched by Spokeo is nearly impossible. First, identifying all of the contributing sites is a Sisyphean task; and second, you create more data each day that you use the web.

As deep-web technology improves, many of these problems will on be exacerbated. Search engines themselves will start acting more like Spokeo: they will aggregate all public information available on any topic—whether a concept, historical fact, or an individual—and return that information to the searcher.

iii. Google and Google Profile

157 Privacy, Spokeo, http://www.spokeo.com/privacy (last visited ___).
As the leading search engine online in the United States, Google makes more information available to U.S. users than any other platform on the Internet. At the same time, Google collects as much information about its users: recall the description of Google Dashboard and the data that Google collects and analyzed on each user. Google, however, makes it a priority to share this information with users—there’s the sense that its proud of the data it collects and analyzes and thinks that information will lead to happier users in the long run.

Google is about empowering users in other ways too: it wants users to actively construct their online persona. Such is the purpose of Google Profile. Through this feature, Google asks users to fill out various information about themselves—information that Google likely already knows from the data it collects elsewhere. The Google profile mimics a Facebook profile in many ways: it asks you for the places that you have lived, your employment and education histories, your relationship status, and fun tid-bits about yourself (which it calls “bragging rights”). There’s another space for links with the question “[w]here are you on the web?” Another tab allows you link up your profile to your Picasa albums and pictures, and another allows you to sets the defaults for your Google Buzz profile.

161 See Part III.B supra.
162 Google Profile, Google, https://profiles.google.com/. You must click on the “Edit Profile” button to see the categories and associated questions.
163 Id.
Google’s description of its profile makes its purpose clear: “Your profile is the way you present yourself on Google products and across the web. With your profile, you can manage the information—such as your bio, contact details, and links to other sites about you or created by you—that people see.” In some ways, Google is trying to get the user to bless what it already knows: it wants to reveal your network and online self. At the same time, Google is disarmingly honest about the publicity and wants the user to think of the Profile as an opportunity to shape actively her online personality. This is a chance for the user to actively and consciously influence what people see when they search for her online. It also gives the user a chance to clarify or correct what she perceives as misinformation, something that Google makes official in its privacy policy. Google of course respects a user’s decision to limit her disclosure, and there is a clear way to delete your profile entirely and shut off Buzz permanently when you edit your profile.

None of this is to say that Google is flawless when it comes to privacy. In fact, Google’s launch of Buzz—where it defaulted to mass disclosure—can be seen as one of the great privacy failures ever. But Google’s profile feature highlights certain trends in peer surveillance. The profile is undoubtedly another vehicle for surveillance, but, like Facebook, it asks users to partake willingly. Yet, unlike Facebook, it doesn’t seek to

165 Google’s Privacy Police, Google, http://www.google.com/intl/en/privacy/privacy-policy.html (“When you use Google services, we make good faith efforts to provide you with access to your personal information and either to correct this data if it is inaccurate or to delete such data at your request if it is not otherwise required to be retained by law or for legitimate business purposes.”).
167 See Grimmelmann, Privacy as Product Safety, supra note __, at 825 (describing Buzz’s launch, including the rapidity with which a lawsuit was filed).
establish social connections; rather, it’s purpose is to facilitate the spread of accurate information, giving a very real voice to the individual.

Furthermore, we are hard-pressed to see any privacy violations in Google Profile, even though it performs each action in Solove’s taxonomy. It increases our ability to surveil; it aggregates information and ties it to a particular user; that information can be used by others and disclosed without our knowledge; and it certainly increases access to that information. But there is something in the way that the purpose of the profile is phrased—“[it’s] the way you present yourself on Google products and across the web”—that makes it all seem less violative. This shouldn’t be that surprising, since, as Solove notes, privacy violations are context dependent. Google has simply clarified that context.

We’ve now seen three examples of different peer technologies that present different privacy problems online. We’ve also seen how these technologies deliver real benefits to users and how the technologies are rarely themselves to blame for the privacy violations. Next, we turn to the common legal and structural solutions proposed by scholars to see if these technologies can and should be regulated on their own terms.

iv. The Problem with Legal and Structural Solutions

Confronted with increased privacy violations as technology (and particularly the Internet) evolves, privacy scholars have focused on a number of potential regulatory and technical solutions. A large number of these solutions treat problems that are not presented here. For example, many scholars have written on data insecurity, whose focus is often database breaches and identity theft. Here, we are concerned with “peer-produced privacy violations” that exist despite the fact that the information was properly
obtained. On Facebook, “[t]he privacy violations are bottom-up; they emerge spontaneously from the natural interactions of users with different tastes, goals, and expectations.” On Spokeo, the privacy problems are merely a function of technology’s ability to aggregate—and we are hard-pressed to say that Spokeo is doing anything wrong by combining and presenting that information. Legal and technical solutions to these privacy violations are notoriously difficult to find, particularly given the fact that each technology offers certain benefits that we would rather not destroy entirely.

Still, a number of solutions have been proposed. Take Facebook—some suggest technical enhancements that would better protect privacy. These include Privacy Enhancing Technology as well as a set of technical controls that allow the user to protect her information. As the reader may recall, however, Facebook already offers a prolific number of privacy options. The problem, as indicated above, is that privacy controls get in the way of Facebook’s purpose, which is to foster social connections. Furthermore, social relationships are complicated and privacy settings simply cannot cope with that complexity. As Grimmelmann observes, a user can decide to share one piece of information with one friend and not another, but that user can’t stop the two friends from getting together and discussing it. “As long as there are social nuances that aren’t captured in the rules of the network (i.e., always), the network will be unable to prevent them from sparking privacy blowups.” Whereas there may be some room for defaults

---

168 Grimmelmann, Saving Facebook, supra note __, at 1188.
169 Id. at 1188-89.
170 Id. at 1184-85.
171 Id. at 1186.
172 Id. at 1186.
to prevent certain types of privacy violations, they cannot prevent every time of privacy blow-up.

Spokeo too has certain technical privacy controls, allowing individuals to remove their profiles. Unfortunately, we have already seen the difficulty in playing that game—Spokeo is just one of many personal information aggregators. What is more, default rules are less likely to work in Spokeo’s case, since the only the “choice” is whether to include each individual by default. In theory, we could require that Spokeo run an opt-in system; that is, individuals would only be listed if they decided affirmatively that they wanted to be listed. Such regulation would not only run into First Amendment problems, it would be bad policy. Search engines would be put on notice—information isn’t public unless it says it is.

Just as with technical regulation, legal solutions often seem to fail to accommodate the complexity of what is happening online. Many scholars suggest that individuals should have a property right in their data and that their ownership of the information would give them the control needed to protect their privacy interests. Yet, too much of the information given on these platforms is owned by multiple users. Who owns a bit of gossip involving multiple Facebook users? Furthermore, ownership would do little to solve the Spokeo aggregation problem, where Spokeo simply aggregates information that the user mostly wants online. Ownership, after all, presumably allows

---

173 Id. at 1200 ("Breakups should be opt-in, not opt-out.").
174 Since Spokeo already combs just the public web, it is unclear what other, more limited searches we could insist Spokeo perform.
175 See, e.g., CITE.
176 See also Grimmelmann, supra note __, at 1194 (discussing the ownership paradigm in the context of data portability).
the owner to grant public access to the information—how could it justify giving public access in some cases but not others? The ownership model simply does help in a situation where people have volunteered information for the very purpose that causes the privacy harm—that they want to share this information with a public audience.

Another popular framework is the tort or products liability analogy. Lior Strahilevitz proposed a multi-factor test to guide courts in their factual inquiry into whether a user had a reasonable expectation of privacy in a suit for public disclosure of private facts.\(^{177}\) By examining social network theory, Strahilevitz concludes that a court could determine with some measure of accuracy when it would have been reasonable for an individual to expect a certain degree of privacy.\(^{178}\) Thus, Strahilevitz concludes that “[w]here a defendant's disclosure materially alters the flow of otherwise obscure information through a social network, such that what would have otherwise remained obscure becomes widely known, the defendant should be liable for public disclosure of private facts.”\(^{179}\)

Strahilevitz did not limit his findings to online social networks, but, as Grimmelmann notes, “[s]ocial network sites—where the social network itself is made visible—are a particularly appropriate place for the kind of analysis Strahilevitz


\(^{178}\) *Id.* at 988. Among other critical factors, Strahilevitz identifies “the prevalence of supernodes in a network and the social distance from those supernodes to the periphery” as two of the main determinants in whether information will spread beyond an initial disclosure. *Id.* at 958.

\(^{179}\) *Id.*
recommends.” The problem, however, with this inquiry is that it puts an unreasonable burden on the defendant, who is in no position to gauge the viability of the initial social network. Straihilevitz envisions tort law acting as “a form of social insurance,” and he has a news media member as the typical defendant. Online, however, we are typically talking about other users, friends even—is it reasonable to ask them to predict whether this information would be disclosed without their sharing it? Perhaps more importantly, is it fair to ask them to expect that their subsequent disclosure will lead to a far wider and more embarrassing disclosure? What happens if, as is often the case, no one person can be clearly identified as responsible for the exposure? Are we to hold each defendant jointly and severally liable? What if the harm wouldn’t have been possible without the actions of thousands, perhaps millions? What if the harm itself is the fact that millions forwarded on the information? These questions doom Straihilevitz’s multi-factor test in the context of online surveillance and interaction. The result would be a system of liability that was unpredictable and fickle and could threaten the good speech and sharing that takes place online every day.

Grimmelmann has proposed a product liability analogy that again focuses on a user’s expectation of privacy. In so doing, he shifts the inquiry away from the individual “supernode” that may be responsible for disclosure and to the design of the site itself. Thus, this is a way to hold Facebook responsible for the ultimate privacy invasion. Grimmelmann starts by noting that users do indeed care about their privacy, but they have a tough time “understanding what will happen to their personal information once

---

180 Grimmelmann, Saving Facebook, supra note __, at 1196.
181 Straihilevitz, supra note __, at 927.
182 Grimmelmann, Privacy as Product Safety, supra note __.
they post it.” 183 In certain situations, he says, users have expectations of privacy, and “[a]t some point, this expectation starts to create its own reality—to become the kind of expectation that creates enforceable duties at law.” 184 He draws the products liability analogy because he believes that social networks deliver those kinds of expectations of privacy; thus, “when [privacy] harms are preventable with better design choices or more careful programming, it makes sense to ask whether the site operator should be held accountable for them.”

The problem, however, is in defining a better design. In the real world of products liability, when a product misbehaves, it is relatively easy to observe it. Grimmelmann gives the example of a hammer that malfunctions. 185 Let’s say that the hammer’s head is poorly fastened. This design is always a liability because it can only lead to an injury, whether actual or reasonably probable. Facebook, however, must be designed both to protect privacy and to facilitate the spread of speech. In fact, it is probably this latter function that is more important. Unfortunately, a design that works in one instance frustrates the user in the other. Of course, this is not to say that there are not good default rules for specific contexts that a site should probably utilize, but to hold them liable for failing to do so when part of the purpose of their design is to spread information easily seems too much. 186

183 Id. at 796.
184 Id. at 805.
185 Id. at 813.
186 To be fair, Grimmelmann doesn’t suggest that the current doctrine of products liability should be directly ported to the online social network context, but he does advocate for legal liability when the design fails to meet certain standards.
I’ve focused on Facebook and social networks in these tort discussions because a few scholars have focused on social networks in particular. Similarly, however, the tort model fails to capture the privacy harm perpetrated by a Spokeo. Straihilevitz’s model is too strict for the same reasons as described above: a user who comes to Spokeo has no idea what the individual’s social network looks like and whether her subsequent disclosure will lead to a privacy harm. It seems equally strange to hold Spokeo itself liable for publicizing the private facts when it (a) finds the information on the public web and (b) does nothing itself to popularize the information (the person searching must still find the information and then share it).

The products liability model, on the other hand, seemingly could not touch Spokeo itself. Spokeo doesn’t make any representations to users and it doesn’t solicit information from users directly. Thus, the products liability suit would again be brought against a Facebook (or other sites contributing information to the public web) for letting private information be public in the first place; Spokeo might provide evidence of disclosure, but it would not be a party to the litigation since it isn’t disclosing any information. What is more, Spokeo aggregates information from many more sites than social networks, including public records provided by the government. As a result, the tort model would provide little relief to a victim of mass aggregation.

The over-arching problem in all of these models is that the same features that lead to privacy violations in some circumstances are features of those platforms as well. These platforms are very good at facilitating speech and interaction, which is often what Internet users want. Essentially, it’s extremely hard to tell when the system is functioning properly and when it isn’t; it only becomes clear after the fact when someone feels that
their privacy has been violated, and often times it is unreasonable (or unconstitutional) to hold the violator or the platform responsible.

Is there anything we should do in terms of regulation? Perhaps, though it has little to do with privacy. Much like the surveillance state, we can make sure that individuals have access to their own information and what is out there on them. Facebook and Google do a good job of this already, although other enhancements could be made. For example, Facebook should offer an privacy feature where you can “untag” photos of yourself that other users post, but those photos should still appear under “photos of you” when you log-in to your profile. That way, a user is able to both untag and keep track of all photos of her out there. Similarly, Spokeo should offer their full report for free to individuals searching themselves. None of these suggestions have to do with privacy. Instead they focus on what Google noticed with its profile feature—that users want control over their information, and the first step to control is to understand what information about you is out in the ether.

In this section, I’ve tried to show why direct legal regulation and rules fail to accommodate the complexity of online privacy harms. I’ve focused on a few of the most promising theories for liability, but other more traditional structures (like contract) clearly won’t work to regulate peer violations. This discussion has built upon the problems we observed in trying to impose a coherent legal regime on the surveillance state and on corporate surveillance. Given we face concrete problems in regulating each of these areas, we will now turn to an analysis of the surveillance ecology on the whole. Of course, in the end it may be the case that we opt for regulation in some of these contexts—perhaps because we overcome the practical hurdles identified with regards to
regulating the surveillance state, or because we decide that the problem of peer-produced privacy harms is just too great. But before we make that choice, we should see what self-help remedies the surveillance ecology offers.

IV. THE SURVEILLANCE ECOLOGY ANALYZED

As I hope I have show, information collection is the mechanism by which our surveillance society functions. Each actor—government, corporations, and individuals—seek out information with the goal of interpreting, analyzing and acting based on what they learn. In an important way, then, there is no big choice to be made between privacy and freedom of information; we, as a society, have already chosen a culture of surveillance. Government runs on information; corporations run on information; we run on information. Because of this, I believe that it is too late to put the cat back in the bag, so to speak.

The government can be seen as starting the transition to the modern surveillance society (although corporations have been surveiling their customers for a long time themselves). What is important, however, is that it is the government’s transition to the surveillance state that makes it impossible to “go back” to an era of rich privacy protection. As Professor Balkin argues, the surveillance state is a permanent form of government. The logic is simple: there are certain events—such as terrorist attacks—that a government will always believe is worth stopping, even at extremely high costs. Frankly, surveillance and information analysis is not that a high of a cost, politically speaking. What is more, there is already strong evidence that regulation will not stop the government (and the executive in particular) from invading our privacy in order to
predict the future and track down bad guys. For example, FISA and the FISA court did not stop the Bush administration from illegally surveiling Americans for years. In fact, all that happened was a post-conduct ratification.

Given this reality—that the government, at least, will always engage in the type of surveillance described above—we need to ask what best protects us from further encroachment. The instinct of privacy advocates is to try to save privacy wherever we can—to regulate corporate collection and retention of data, to expand tort liability, and more. Such comprehensive legal expansion would not only have bad economic and social consequences; it would also actually undermine our most effective mechanism of accountability of the government—mutual surveillance. Instead of throwing up barriers, we should strive to have people participate fully in the surveillance ecology.

Participation occurs along two lines: first, we can all surveil others in the system; and second, we can contribute personal information about ourselves to the system. Both are essential to creating an accountable surveillance society. On the first, we—and corporations—need to be information consumers in a very broad sense. We need to consume information not just about ourselves but about the government as well. Again, Google provides a good example—their Transparency Report records and displays by location the number of requests for user data that they get from governments across the world. As Google Vice President of Public Affairs Rachel Whetstone said when the tool was launched, “Increased government censorship of the Web is undoubtedly driven by the fact that record numbers of people now have access to the Internet, and that they

---

are creating more content than ever before.” The tool provides one way to understand how many and what types of requests the government makes on a regular basis.

Other accountability measures can occur on a much more micro level. One way to combat the threat of bureaucratic harms that have infused our discussion so far is to ensure that individuals surveil the bureaucratic institutions themselves. As we interact with the government, we should make a record of it—contribute it to the surveillance ecology. Publishing has never been easier, as has been demonstrated again and again in this paper. People can use their Facebook accounts, their blogs, their RSS feeds, etc. to share stories about our government. But it doesn’t stop there. These same platforms give us an incredibly convenient way to aggregate this information, make connections, and change public perception. All of this depends on a culture of mutual surveillance—we need to follow each other so we know what is happening.

Contribution and participation in the surveillance ecology achieves another purpose—it allows us to correct wrong information about us. For example, Facebook and Google Profile allow a user to shape her own reputation and correct misinformation. These private tools of surveillance afford users new opportunities for self-expression and self-creation. Instead of merely being the subjects of information collection, we are now active participants. Although these new modes of authorship are not the perfect substitute for the loss in privacy, it is significant that the Internet gives us new ways of participating in society as autonomous agents. As one scholar puts it, “not every diminution of privacy

---

necessarily harms autonomy, or any other valuable interest. Some losses of privacy may be neutral, or even positively beneficial to other values, including autonomy.”¹⁸⁹ To be clear, I am not arguing that we have gained more than we have lost—just that we have gained as we have lost.

None of this is to say that we won’t shape privacy at the margins. Rather, the argument is that we don’t necessarily want comprehensive regulation because that would counter-intuitively give the surveillance state more power. In a closed off, un-transparent society, the government can break its own rules because there is no one to hold it accountable. As Richard Nixon famously said in his interview with David Frost, “[W]hen the President does it, that means that it is not illegal.”¹⁹⁰ We can, however, create accountability through information sharing and participation in the surveillance ecology. It is better to have a mutual system of “oversharing” than to have the government monopolize information collection and analysis—and there is no way to keep the government from collecting and processing information.

The interaction between government and the private sector that I have described in this section probably sounds both pessimistic and optimistic: pessimistic insofar as the surveillance state is a permanent structure of government about which we can do nothing; and optimistic in terms of our prospects at creating accountability through mutually assured surveillance. The interaction that I have described above, however, is what we should aim for, and there are ways that we can help cultivate it.

First and foremost, we should ensure that there is diverse ownership of the data inputs. Although it may seem strange, it is better to have many owners of the information because it becomes less valuable the more it is shared. As a result, it is a good thing that corporations engage in surveillance. Second, and as indicated above, we should create tools for analyzing the information that is widely available on the Internet. Technological development is occurring on its own, and new algorithms that can better comb and combine databases will ensure that individuals can use information to draw conclusions much like the government does. Again, corporations can and should play a role here. Third, we should educate people about the surveillance ecology and teach them how to use the Internet to author their own reputation, not just how to obscure or protect oneself from surveillance. Fourth, we should use technology and the law to encourage participation and transparency throughout the system. Although it may seem tangential, the net neutrality debate is essential to equal participation in the surveillance and information ecology. “[I]nfras[tructure] and the way it is engineered matter. It embodies policy choices at the same time as it sets limits on policy choices.” \(^{191}\) One of the basic principles of the Internet since its inception has been that the network is “stupid.” That is, “[t]he network simply forwards or routes the data packets and does not—and cannot by architecture—discriminate or differentiate traffic generated by different applications.” \(^{192}\) This technology guarantees every participant access to the necessary platforms and means of communication to ensure equal participation. Major telecommunication

\(^{191}\) Internet Governance: Infrastructure and Institutions 3 (Lee A. Bygrave & Jon Bing, eds., 2009).
\(^{192}\) Lawrence B. Solum, Models of Internet Governance, in Internet Governance, supra note __, at 62.
companies, however, have asserted their right to discriminate based on content.\textsuperscript{193} Law has a role to play here in ensuring that the net remain neutral.

Of course, none of what I have suggested rids us of the peer privacy harms that we identified above. For better or worse, most of these harms are an inevitable part of modern society—there is nothing we can or should do about Facebook “burns.” It is possible, however, that over time our norms will realign as we understand that every person has individual incidents that are not indicative of her life as a whole. In fact, other aspects of the surveillance ecology may help with this norm-adjustment; for example, aggregation encourages us to see the bigger picture rather than the details (and statistical outliers should in fact be ignored). This, of course, may not be of much comfort to privacy advocates.

Finally, we should ask whether there is anything we can do to limit or improve government surveillance. The answer, perhaps surprisingly, is yes—we can regulate the use of information. Here, I do not mean broad uses, such as “for surveillance purposes.” Instead, we can regulate specific uses of specific information. For example, the government uses certain terms like the word “bomb” in an email to trigger further surveillance. We should limit what categories of terms trigger this follow-up—clearly, the government should not use the name “Mohammed” to trigger increased surveillance. Extrapolating from this, the government should not use religious, racial, or other suspect classes as proxies for wrongdoing. Is the government guaranteed to follow these regulations? No, but such regulation would send an important signal about what our society believes to be morally valid: we will tolerate surveillance, but only if conducted

\textsuperscript{193} See COHEN, supra note __, at 77.
on proper terms. At the same time, in a society where we also surveil the government, it gives us something to point to if we catch them using information improperly.

V. CONCLUSION

The surveillance game is much more complicated than it once was—everyone is involved, whether they think so or not. In the wake of 9/11 and the War on Terror, the government has pushed the modern surveillance state to the front and center of society. Its business is the collection and analysis of information, personal and otherwise. At the same time, corporations have developed new and more nuanced ways to track their own customers, with the hope of using that data to better market products to willing consumers. Their data collection also supports the surveillance state, giving it the necessary informational inputs by which it operates. Finally, individuals have increasingly gotten into peer surveillance. We all have an increasing appetite for information about our friends, family, acquaintances, and even strangers. As one scholar notes,

[w]e claim to respect privacy, but in fact we devour the private secrets of hundreds of people every day. We do value privacy, but not as much as we hunger to know—to know the shocking details of scandal, to see the drama of terror or grief or humiliation, to understand the strangeness of our neighbours. The law merely reflects our own ambivalence—it would like to protect privacy, but the arguments for invading it are usually irresistible.194

---

This peer consumption of personal details and stories has to many only exacerbated the privacy harms of the modern era.

In Part IV, however, I offered a rosier picture of the surveillance society. Given that the surveillance state is a reality of modern governance, we are actually better off in a culture of mutually assured surveillance. Whereas mutual surveillance does not save privacy, it does offer us moments of self-creation, self-determination, and social accountability. Rather than protect privacy in among corporations and individuals, the surveillance society should seek to ensure is three things: first, that information is collected and shared widely; second, that the information is used appropriately (and not subject merely to bureaucratic control); and third, that individuals are able to participate in the culture of surveillance.

In many ways, what I have described is a return to public life. In the modern surveillance state, we are all responsible for holding the government accountable for its actions—and we can no longer expect complete privacy in our own lives. Technology has made privacy increasingly difficult to obtain, but it has also given us the tools to shape our surveillance culture. Now, we need to use them.