Living Books About Life

Ubiquitous Surveillance

Edited by David Parry

By Nachoman-au (A digital photograph taken by myself.) [GFDL (www.gnu.org/copyleft/fdl.html), CC-BY-SA-3.0 (www.creativecommons.org/licenses/by-sa/3.0/) or CC-BY-SA-2.5-2.0-1.0 (www.creativecommons.org/licenses/by-sa/2.5-2.0-1.0)], via Wikimedia Commons

OPEN HUMANITIES PRESS

Living Books About Life

OPEN HUMANITIES PRESS JISC

Ubiquitous Surveillance

edited by David Parry

Introduction

Ubiquitous Mobile, Persistent Surveillance

In 1996 when John Perry Bartlow wrote *A Cyberspace Independence Declaration*, internet pioneers hoped that the online world Bartlow was describing would come to pass. While Bartlow’s rhetoric was admittedly 'grandiose,' his central claim, that the internet was a place of freedom separate from the limits of the physical world, reflected the utopic atmosphere of the time. The technological revolution, in particular the rise of the digital network, seemed to point to a future 'where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity' (Bartlow, 1996). While not everyone in the late 90s could be characterized as a cyberutopian, the dominant mood harbored a sense that the digital network would bring with it newfound, unregulatable freedoms.

Flash forward fifteen years and the present looks significantly different from the one envisioned in the 1990s. Instead of a future of Second Life and virtual worlds, we ended up building one that more closely resembles Minority Report. Advances in technology, an increasingly regulated and monitored digital network, and a general atmosphere of securitization have yielded a world of ubiquitous, if not always visible, surveillance. Consider for a moment how much daily surveillance people are exposed to that would not have been possible fifteen or twenty years ago:

> > In nearly every place inhabited by people one finds video cameras. In the United Kingdom (one of the most recorded populaces in the world) there is one camera for every twelve people, and the average Londoner is caught on camera hundreds of times a day. Most businesses in the United States have some type of video equipment set to monitor both employees and customers. Public spaces are now recorded, and many of their cameras have a live feed available for remote viewing by anyone with an internet connection. Further, with the growth of mobile computing devices (smart phones) many people now carry video cameras with them everywhere. With the quality of cameras increasing, and their price decreasing, the trend is likely to continue, with little of our daily lives not being recorded by some video device.
>
> > Data mining is in. Our online activities are also increasingly monitored, which produces extensive data trails. Corporations wish to monitor every website people visit in order to more effectively market and advertise their products and services, while governments move towards gaining the power to wiretap nearly all internet traffic. Search engines record the history of submissions, painting a detailed picture of a user's life, while social networking sites get users to record and publicize their offline lives. In countries such as the United States and the United

> Kingdom, government-backed surveillance projects such as ECHELON are reportedly capable of intercepting all forms of data traffic, from faxes and telephone calls to email and internet traffic.
>
> > As if the records of our daily lives were not enough, most of this information is being recorded and stored in massive databases. While science has entered the era of big data, public policy researchers are also now realizing the power of massive data collections. Demographic data of nearly every part of our daily lives is recorded and stored for analysis. This data is used not only to surveil existing populaces, but to help predict trends and future developments, monitoring the future before it even happens. And while this type of research offers serious rewards—for example, helping to curb the spread of disease—it also brings with it serious new social concerns.

It might seem odd to include the concept of 'surveillance' within a book series about life, but one claim I would like to advance as part of this collection is the idea that one cannot consider life without also considering the data that life produces. Indeed, in this contemporary data age, the definition of life is intimately tied up with the ability to produce data. I would argue that life is now defined by the ability to produce data. The way one testifies to being alive, testifies to life, is through the production of data, from the very literal data produced by health monitoring (EKG meters, pulse, cholesterol level) to the constant stream of life data that is produced and shared online (disappearing from Facebook for a few days often elicits questions such as 'Are you still alive?'). If something is not online it probably does not exist ('Can you Google it?'). If you lack a data trail you might have a body, but you do not count as being alive ('I am sorry; I can't help you. I don't have a record of you in our system').

Within the humanities when thinking through the complex problems raised by the digital network and ubiquitous surveillance one of the first places theorists turn is to Foucault's analysis of Bentham's panopticon, the prison that makes its inhabitants feel as if they are always being monitored, to the extent that they begin to self-monitor and remove the need for heavy discipline and punishment. Foucault argues that discipline and punishment becomes internalized when surveillance is a constant possibility, and the ever-present citizen surveillance of our current age seems easy to fit into his model. Indeed, in much of the theorization of modern-day surveillance, the information age is often referred as the age of the always-on-panopticon, and Foucault is accordingly the beginning and end point of these discussions. But there is a disconnect between this type of approach to understanding surveillance and what is actually occurring within the research fields. As these collected articles demonstrate the surveillance system being constructed far exceeds the panopticon, producing a social space not so easily mapped onto the one Foucault imagines.

As Siva Vaidhyanathan has argued, the panopticon is perhaps a theory not so easily applied to the current state of affairs. While the panopticon works precisely on the condition that you know you are being watched and thus alter your behavior, contemporary surveillance often operates on the condition that you do not know you are being watched. It is our lack of awareness of the extent to which we are surveilled that often serves as one of the strengths of the system. Web monitoring, cameras and data collection all work by recording and analyzing 'natural' behavior. The more one acts as if one is not being monitored, the more useful the data is. Sure, people generally know that information that they post online is observable by a wide range of individuals, but few are aware of the extent to which their lives are monitored, observed, and subsequently controlled outside of the arena of social networking. A close reading of these articles shows how researchers are busy constructing (often consciously) a ubiquitous surveillance system, one which operates far beyond any awareness of the individuals being monitored. In fact, an individual's choice to not post information online might provide merely a false sense of privacy, since whether you share or not, your life has been recorded.

However, if Foucault’s panopticon is right about one aspect of our contemporary lives, it is in the conception that the real power is not with Big Brother, but rather distributed throughout the social space. In Foucault’s account the state observes people and thereby produces altered behavior, but also, importantly, a common sense of correct behavior develops, and individuals alter their behavior as well in the name of social conformity. In this sense, Big Brother is not the government—rather, Big Brother is us. The ease with which we can monitor each other and self-monitor our behavior accordingly resembles this particular aspect of Foucault’s panopticon on steroids.

There is much to be said about our surveillance society, and perspectives from a range of disciplines and foci are crucial to understanding this phenomenon. It would be impossible to adequately cover this whole field in an edited collection, even one not necessarily limited by print production. Hence I have chosen to group this collection into four sections, each dealing with one specific area of surveillance. In many cases the articles, in keeping with the turn in science to 'big data', are heavy on math, analyzing large data sets. But in each case the particulars of the math are not as important as the overall picture the research paints. Engaging the nuances of the calculus involved is not required for understanding the general theme of the articles (to be honest, in many cases I do not understand all the math being used). Understanding the contours, direction, and possibilities of this type of research, though, is important.

In the section and article summaries below, I have tried to highlight what I see as the important issues in each. Surprisingly, few of the pieces recognize a problem with the technologies or policies being developed. While a few mention in passing privacy concerns, many do not, and fewer still (even in the legal section) recognize how technological advancements are bringing about massive disruptions in the way we conceive of public and private realms. Yet, none of the articles seem to recognize the complexity of the problem. A new equilibrium between public and private is likely to require re-negotiation through legal, technical, and cultural means—no single means being sufficient. However, the first step in any such re-negotiation is understanding the depth and contours of the problem. I offer the following collection of articles as one step in understanding our contemporary predicament.

So, read on, and try not to get too paranoid.

Note: There are actually two tables of contents for this book. The first is a standard (able-to-be-printed) one, and the second is an interactive map containing article summaries, which is geo-located with surveillance cameras.

Knowing Everything: Data Mining

The first of the sections looks not only at the degree to which data is now being mined, but also at what can be done once that data is harvested. In 'All Liaisons are Dangerous When All Your Friends are Known to Us,' Daniel Gayo-Avello demonstrates the ease with which private information about an individual can be gleaned, not from things they write, post, or share, but rather by simply looking at the contours of an individual's network connections. Perhaps the most publicized example of this type of analysis was 'Project Gaydar,' a software program created by MIT graduate students that scanned individuals friends and lists and was able to determine through association an individual's sexual preference, irrespective of whether or not said individual indicated it as part of his/her profile or public discourse.

In his study, Daniel Gayo-Avello shows how, relying on the principle of homophily and analyzing relationships, in this case on Twitter, one can determine sex, age, religious or political affiliation, race or ethnicity, and sexual orientation of a given individual with a relatively high degree of precision. This is particularly troubling, as Gayo-Avello points out, because each of these categories represents a class of information individuals might want to protect for fear of discrimination. While previously researchers had demonstrated the possibility of determining information about particular individuals via their relationships, Gayo-Avello tests a new algorithm which demonstrates the ability to precisely garner information while knowing information about only 1% of users. In other words, while you and your friends might keep information guarded and secret, data miners only need to know information about 1% of users to infer information about the remaining 99%.

This type of surveillance or threat to privacy is particularly nefarious, for it exploits the very thing that makes social networks useful socialization tools. The more an individual connects, the more it is possible to glean information about that individual. Once connections are known information is also known, and hiding connections is not a viable choice as the publicness of these connections is what creates their usefulness. While in the end the algorithm Gayo-Avello develops works with greater or lesser precision depending on the category, the key here, as he points out, is not the efficiency, precision, or accuracy of the algorithm but rather the realization that these types of data analysis will become increasingly more efficient, precise, and accurate.

While Gayo-Avello gives us a way to discover undisclosed information about individuals by analyzing their network of friends, the authors of 'Googling Social Interactions: Web Search Engine Based Social Network Construction' develop a way to discover who your friends are, even if explicit links on social networking sites have not been constructed. As the authors outline in their abstract, 'the exploding amount of automatically generated data has completely changed the pattern of research' (Lee et al, 2010: 1). By leveraging the large amounts of data already generated, a wide range of fields are able to conduct research previously not possible, with very little investment. Often, it is no longer necessary to conduct experiments in order to generate data for analysis; instead, one need only analyze existing data sets.

While this type of data analysis can lead to many positive social outcomes (see the third section on health, for instance), the implications for privacy and surveillance are, to understate the case, significant. In this particular instance, the authors demonstrate how it is possible simply by analyzing Google to discover and map social connections. In short, Google knows who your friends and rivals are, and knows the size and range of your social network, without any individual explicitly creating links. To demonstrate how this works, the researchers in this paper focus on the 109th Senate of the United States, and by leveraging Google search results are able to map the social network of these individuals.

Several things are worth noting here. First, as already mentioned, this type of research did not require the generation of any new data; that is, individuals were not questioned or surveilled in order to create the social network maps, because the researchers relied on the 'tremendous amount of data which can be useful' that was already in existence (Lee et al, 2010: 9). Second, this type of research is incredibly cheap. Third, although some of the initial conclusions seem rather obvious—Democrats are more closely related to each other than to Republicans (and vice-versa)—by repurposing the data new patterns and interactions develop. For example, by adding in another layer to the data, the researchers were able to not only map social interactions between Senators, but also between corporations and the Senators. Fourth, this type of analysis is dynamic. That is, not only can one look at the social graph at any particular moment, but it is also possible to look at the change in the social graph over time.

And finally, and perhaps most importantly: this is not just about political figures. While it might seem (indeed I would argue it is) civically useful for citizens to have access to this type of data analysis, the researchers just chose the Senators as a representative sample. In the future, as more data is produced and analyzation techniques continue to improve, it will be possible to perform this type of social graph analysis for any group of individuals, not just prominent public figures; everything from workplace surveillance (determining which employees are the most connected) to neighborhood social graphs (determining which individuals are closest friends). It will become possible to generate graphs for any individual, showing his or her change in friendships over time, in addition to his or her relationships to corporate and public institutions (i.e. I used to be more connected to Walmart, but now I am more of a Target person).

One of the primary challenges for those seeking to analyze data is not the collection of data, but rather the reverse, that too much data has been collected. Traditionally, data has been useful only after it has been collected—that is, the data is generated and stored, and only once stored is it then analyzed for knowledge. But due to the high volume of data being generated now, as Kholghi and Keybanpour point out, it is 'impossible to store an entire data stream or scan through it multiple times' (2011: 2508). As a result, researchers are looking for a way to analyze the data as it is being generated, which would prove to be a far more powerful method of surveillance.

In their article, Khloghi and Keyvanpour look at a range of approaches to solving this problem for researchers. Whether or not their particular analytic framework is adopted is in the end far less significant than the future towards which they point. Although recent political discussions (especially in Europe and the United States) have begun to express concern with the level of data mining taking place, we are only at the tip of the proverbial iceberg of data, analyzing only a small subset of all the data being produced. Ultimately engineers will figure out a way to analyze the data as it is being generated, exponentially multiplying the range and power of surveillance, whether by private companies or government agencies. As with many of the articles in this collection, Kholghi and Keyvanpour express little concern over the social implications and challenges to such data analysis, instead treating this as merely a technical problem in search of a technical solution.

In the interest of saving the most disturbing article for last, 'A Survey of Deep Packet Inspection for Intrusion Detection System' analyzes systems which monitor not only network traffic, but the contents of packets shipped on the network. If you think of the internet as a mail system (an admittedly imperfect analogy), and the information sent via the internet as closed envelopes with data inside, currently it is rather easy to read information on the outside of the envelope (the address, the return address, etc.), but more difficult to read what is inside the envelope (the contents—you would have to steam open the envelope, for example).

Deep packet inspection allows network monitors to not only read the packet information (what is on the outside of the envelope) but to monitor its contents (what is on the inside). At the most nefarious level, deep packet inspection would not only allow for robust filtering, but enable a network intermediary to alter the contents of the packet and deliver it without the recipient becoming aware of any change.

Deep packet inspection is currently limited, for it requires resource intensive computing, both on the hardware and software sides. In this article, the authors outline the 'challenges and goals' to developing deep packet inspection (AbuHmed et al, 2008: 1). Tellingly, deep packet inspection is treated as a positive (no negative consequences are mentioned), as the authors highlight the ways deep packet inspection can be used for network security (for example filtering spam), with no recognition that these techniques pose a serious risk. Deep packet inspection is treated both as an inevitability and a technical problem to be overcome through faster computers and better algorithms. However, it is because of the power of this type of inspection that many internet advocacy organizations have cautioned against developing and implementing these types of technologies. Deep packet inspection would enable private or government intermediaries to not only monitor but regulate internet traffic. Not surprisingly, one of the most famous cases of deep packet inspection is its use by the Iranian government during the 2008 uprisings. Using technology purchased from Siemens and Nokia, the Iranian government used deep packet inspection to block and monitor certain kinds of traffic. At the time this technology was relatively new and thus its implementation was limited by the available hardware and software, but as this article demonstrates, those limits are soon to be overcome, making possible a new level of surveillance of internet traffic.

Somebody is Watching You: Cameras Everywhere

Often the most invoked image of the modern surveillance society is the camera: eyes watching everywhere. As the authors demonstrate in 'Motion Pattern Extraction and Event Detection for Automatic Visual Surveillance,' video-based surveillance will soon be far more powerful. Currently, the limiting factor in video surveillance is the human component. That is, regardless of how many cameras one places for surveillance someone still needs to look at all the footage to ascertain its significance. Despite all of the CCTVs in London, someone still has to view all of the recorded footage, or view all of the cameras live, and a viewer can only watch so much footage or monitor so many cameras at once (a fact the UK tried to account for by encouraging citizens to watch CCTV footage at home, turning the populance into a crowd-sourced group of video monitors). But, researchers are working to remove this limitation using computer algorithms to process and monitor video footage as it is recorded. These automated surveillance systems are designed to 'integrate real-time and efficient computer vision algorithms in order to assist human operators' (Benabbas et al, 2011: 1).

In this article, the authors outline an algorithm which can analyze video surveillance footage to determine six different crowd-related events: walking, running, splitting, merging, local dispersion, and evacuation. While still in its infancy, this type of computer-aided analysis already proves to be particularly accurate in analyzing a range of crowd behaviors. By focusing on 'groups of people rather than individuals' the algorithms are able to detect and predict the pre-targeted events (Benabbas et al, 2011: 6). This type of surveillance is then applied to a range of scenarios, including urban populations with cars and people, and low and high density areas. As the authors point out, future research is likely to improve results and performance of this method. Most importantly, however, in the closing sections of this article, the authors point towards the desired future in which computer aided systems will be able to track 'single persons and [detect] abnormal behaviors' (Benabbas et al, 2011: 14).

The second article in this section, 'A Logic Programming Approach to Behavior Recognition,' describes one of the particular paths to developing a computer surveillance systems which can detect and recognize individual human behavior (as opposed to large group analysis). The goal for these researchers was to use an automated system not to detect short term behaviors, but rather to detect long term ones (a far more difficult task). By using the computer to detect chains of short-term behaviors these researchers were able to have a computer detect long-term ones with some success. The system works by detecting 'short term behaviors that, if satisfied, lead to the recognition of long term ones' (Artikis & Paliouras, 2009: 1).

It is worth recognizing, with regard to this article, that the technique the researchers used, called Event Calculus, allows for computer-based reasoning about events over time rather than a static state. In other words, the computer using the algorithm is not limited to static images but can analyze and 'reason' about changes in images over time. If the authors of this article are correct, their techniques point the direction to an efficient way for computers to monitor, in real time, video footage for specific events. The goal is ultimately to 'teach' the computer a series of short term behaviors which it could then use to predict more complicated ones, making behavioral surveillance possible. This would enable extensive automation of surveillance footage: computers could analyze a data stream and alert a human user to focus on a given camera when an event threshold is reached.

One tenet of technological development is that what is initially expensive and available only to governments and large institutions soon becomes widely available for personal use. In 'GPRS Video Streaming Surveillance,' Pushpavathi, Selvarani and Kumar describe a system based on existing technologies for personal video surveillance. The authors demonstrate a system which, by leveraging already widely available technologies, would produce high quality images yet meet low bandwidth requirements, making it possible to view surveillance footage via a 'mobile phone from a remote location' (2010: 40). The system as designed is presented through a simple interface so that 'people can use it with the utmost ease' (40).

As with other articles in this collection the particulars of the researchers’ claims are not as important as the general direction that this research represents: in this case, making surveillance ubiquitously and easily available. Imagine being able to tune into surveillance cameras placed at your home while you are away from your house by simply pulling out your mobile device, even if you are in a low bandwidth area, and receive text message updates about possible intrusions. But unlike many of the other authors in this collection, Pushpavathi, Selvarani and Kumar recognize the dual nature of this technology which not only provides 'powerful opportunities for increased independence and a higher quality of living for inhabitants . . .they also pose threats, regarding security problems' (40). Not only will governments and large institutions posses the technology of surveillance (Big Brother is watching) but individuals will be monitoring as well (everybody is watching everybody).

Again, saving the most disturbing article until last, the authors of 'SwarMAV: A Swarm of Miniature Aerial Vehicles' describe their research into building not just surveillance cameras, but cameras attached to tiny aerial vehicles: vehicles which can coordinate with each other and operate semi-autonomously. In the future, robotic insects will have cameras and they will be recording your every move.

As the authors indicate, a swarm of tiny flying cameras posses several advantages (or disadvantages, depending on your take). These unmanned flying cameras would be multiple, thus offering a high level of overlap. If one camera fails, there will be others in the area also capturing the same footage. Further, a swarm of cameras could exchange information with one another and coordinate to monitor dynamic situations in a way that individual and fixed cameras cannot. It is clear from this article that much in terms of the technology needs to be developed. However, as micro electronics, cluster computing, and robotics develop further, this technology, like all the others mentioned in the collection, will only improve.

Monitoring Bodies: Surveilling Health

It is not surprising that the researchers focusing on using surveillance to produce social good are also the ones who recognize the drawbacks and substantial concerns associated with creating a ubiquitous monitoring. Perhaps nowhere is this problem as heavily discussed as it is within the public health field, where increased surveillance could help to better understand disease, more efficiently allocate resources, and monitor epidemics, but where individuals' privacy is also a primary concern. In 'Should Data from Demographic Surveillance Systems Be Made More Widely Available to Researchers', the authors point out the limits to sharing health surveillance data.

The limits that these researchers highlight are not primarily concerned with privacy. Rather, technical or financial limitations are noted, along with the lack of coordination between organizations. These authors argue that given the correct circumstances, researchers are more than willing to share data without concern for organizational or institutional ownership. The authors call for a wider sharing of public health data amongst organizations. But as this data becomes more widely available and easily distributed, concerns over exploitation will also increase. The Group Insurance Commission in Massachusetts serves as a cautionary tale: after publishing anonymized data about claims, researchers at MIT were able to re-identify patients, linking medical histories to particular individuals, famously retrieving the medical records of then Governor William Weld.

The second article in this section, 'Web GIS and Public Health' looks concretely at sharing one of the types of public health data: geospatial data. As the authors point out, geospatial data 'provides new opportunities to advance disease surveillance, control, and prevention, and insure public access and community empowerment in public health' (Najafabadi & Pourhassan, 2010: 1). But as the authors quickly caution, geospatial data is particularly sensitive and its usefulness also makes it a particularly rich target for exploitation. Even if the data is anonymized to obscure the exact location and only provide a general identifier such as a zip code, the risk of de-anonymization is high. Coupled with gender and date of birth, a zip code can be used to uniquely identify 87% of the US population.

But this does not mean that geospatial health data should not be collected, for as the authors argue, it can be uniquely powerful in preventing disease outbreak and empowering local communities. What is more, for the most part this type of data has already been recorded and stored for use; it is simply a matter of choosing how and under what circumstances to make it available.

Whereas the first two articles in this section focus on the general ethical concerns facing health surveillance and data collection, the third article in this section brings the controversy surrounding health surveillance into stark focus. In 'Conducting Unlinked Anonymous HIV Surveillance in Developing Countries: Ethical, Epidemiological, and Public Health Concerns,' the authors demonstrate the complexity of ethical concerns facing public health officials dealing with HIV surveillance.

The essential question is: to what degree does the need for 'population-level surveillance' override the concern for individual patients? As the authors note, doctors have typically erred on the side of individual privacy, but with new epidemics concerns have to be reconsidered (Rennie et al, 2009: 32). Specifically, the researchers are interested in looking at how even Unlinked Anonymous Testing (UAT), despite its anonymous and unlinked nature, is still fraught with ethical concerns. In many cases, the way that the UAT is conducted violates the ethical spirit behind UAT. In some cases UAT was performed without patients' consent or in a manner that allows the test subject to be easily identified. Indeed, only 28% of the programs analyzed were found adequate in terms of collecting and protecting the data. Ultimately, health workers face an unsolvable conundrum: Only fuzzy data is safe. Fuzzy data is useless.

In November of 2008, Google launched Google Flu Trends, an attempt to use data generated from search terms submitted to Google to predict Flu outbreaks. The connection here is rather simple: as people begin to show flu-like symptoms they are likely to submit those symptoms as search terms to Google, allowing Google to monitor spikes in certain flu-related search terms. While the accuracy of Google Flu Trends is still up for debate (does it not match CDC data because it is wrong, or because its data is better than the CDC?), the concept nevertheless serves as a jumping off point for other researchers. In the final article in this section, the authors of 'Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance,' demonstrate how internet search queries could serve as possible sources of data for the 'early detection and monitoring' of dengue epidemics.

It is often difficult to collect data on disease outbreaks, either because of lack of infrastructure for accurate collection or because it is difficult to get individuals to self-report. As a result, researchers are looking to other internet-generated data to serve as possible predictors. In this case, the authors used only search terms submitted to Google in an attempt to predict dengue outbreaks in a range of countries: Bolivia, Brazil, India, Indonesia and Singapore. By refining the algorithm and removing noisy data, the authors were able to use Google to accurately predict dengue outbreaks. As they note in the closing paragraphs, despite its usefulness, such a surveillance technique raises privacy concerns. This is a greater concern than the researchers seem to indicate, however, as one realizes that as with many other areas, the de-anonymization of data is going to become increasingly simple, especially as attempts to gain even more powerful tools for passive surveillance branch out beyond merely monitoring Google to monitoring a range of web activity (private emails, Twitter messages, Facebook Posts, web traffic, etc.)

Judging Privacy: Legal Issues

The final section of this book moves away from articles produced in the scientific community and looks at how advances in surveillance technology are being discussed in the legal field. The now ubiquitous distribution of surveillance technology, coupled with the rise of digital networks and social media, means that the courts are having to determine a wide range of developing privacy concerns. How much data is the government allowed to collect and store on any individual without a subpoena? What type of internet traffic is the government allowed to monitor? Is wiretapping the internet the same as wiretapping phones, or does it require a different legal framework? And what of corporations? Should there be limits on how much information corporations can collect on individuals? What limits should be placed on what corporations do with data once they have collected it? The legal questions surrounding the evolving technology landscape are legion.

Any discussion of these legal questions necessarily begins with Warren and Brandeis's famous article, 'The Right to Privacy.' Written in 1890, this article deals with a prior moment of technological transition. Largely responding to the rise in newspaper publications and photography, Warren and Brandeis were concerned that technologies now enabled privacy to be breached in new ways. 'Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life,' such that the technological transition requires the court to 'define anew the exact nature and extent' of the protection the government provides to the rights of its citizens. Thus they argue that despite the lack of an explicit right-to-privacy guarantee in the U.S. Constitution, the courts should recognize that many of the rights enumerated in the Constitution are in effect designed to guarantee privacy, and that given technological advances, it is important for the health of the citizens and the society to directly recognize this right. While clearly focused on an earlier technological shift, there are two important reasons to consider this article anew: one, the Brandeis and Warren article serves as a beginning point in legal discussions about privacy; and two, their approach, that technological transitions require a rethinking of legal values, is one that scholars today ought to consider.

Daniel Solove has become one of the leading legal voices arguing for the need to reconsider privacy in the surveillance soceity. In this article, 'Data Mining and the Security-Liberty Debate,' Solove argues that in a post 9-11 world, civil liberties are often traded for security. Although Solove recognizes the need for balance, he argues that 'there are systematic problems with how the balancing occurs that inflate the importance of the security interests and diminish the value of the liberty interests' (2008: 345).

Accordingly, Solove looks at the range of arguments invoked in the name of security and systematically demonstrates how the legal debate does not often correctly consider all of the tradeoffs. All too often, he argues, the courts now submit to the executive branch when it comes to privacy, accepting 'security' arguments. As Solove is keen to point out, privacy is not merely an individual good, but a social one as well. That is, ubiquitous surveillance not only is a threat to individual liberty but to the social sphere as well, and courts would do well to recognize this. Despite the seductive powers of data mining, we should be careful to not allow the government too wide a purview in using or collecting all the available data.

In the third article in this section, Omer Tene takes up the legal questions surrounding one of the most powerful internet corporations: Google. In 'What Google Knows: Privacy and Internet Search Engines,' Tene argues that despite Google's unofficial policy of 'Do No Evil,' there are substantial reasons for concern, and that Google is evolving into a company with a reputation as a 'privacy menace' (2007: 4) The concerns here are obviously warranted. Search queries alone represent a startlingly rich cache of information on any particular individual. Few would want their full (or even partial) search histories publicly revealed. What is more, search is just the beginning of all the information Google has on any particular individual: email, calendars, social networking, videos, documents, and maps represent just some of the information Google stores on us.

As Tene argues, in a similar manner to Warren and Brandeis, the technology has changed the legal calculus. While in prior moments this type of information might have 'been in the public sphere, it was protected de facto from all but skilled investigators or highly motivated researchers' (2007: 7). The internet, broadly speaking, and Google specifically have changed what used to be by-default private into something that is now by-default public. One has to actively work to protect privacy. Because we are, as many have pointed out, not Google’s customers, but rather their product (we are what they sell to advertisers), new types of legal protections will need to be developed. Tene outlines many of the concerns around Google, and by extension many other internet services, and in the end readers realize that right now there are very few legal solutions to the precarious situation in which we currently find ourselves.

The final essay in the legal section takes up what the author Paul Ohm considers to be the greatest threat to privacy on the internet: The Internet Service Providers. In 'The Rise and Fall of Invasive ISP Surveillance,' Ohm argues that while ISPs have for the most part respected user privacy, in the near future we are likely to see this change, because corporations and governments see them as a way to surveil the internet, whether for security or profit motives. Thus the need for legal intervention to 'distinguish between an ISP’s legitimate needs and mere desires' is greater than ever.

As the carriers of digital information, ISPs are in a unique position to monitor all of our communications. Situating his discussion within the history of privacy law, Ohm demonstrates that existing law, for example legislation covering wiretapping, cannot address the particular concerns associated with ISPs. And although he recognizes the social harms in allowing pervasive surveillance, Ohm insists that the harms to individual liberty alone warrant legal intervention, and should serve as a guiding force in crafting new legislation. While in the early days of the internet, surveillance technology was simply not powerful enough to monitor all its traffic and thus was necessarily limited, advancements have meant that surveillance can now be both automated and targeted in a way to allow widespread monitoring of internet traffic and content. Thus according to Ohm, 'at least in the near term, ISPs will continue to have the advantage . . .a technological constraint that used to protect privacy has since evaporated' (2008: 15).

In the end, Ohm proposes several solutions, a multifaceted pragmatic approach to protecting privacy and legislating against pervasive ISP surveillance. Disturbingly, though, his solutions seem not only unlikely to be adopted, but also not wholly up to the task. They are inadequate not for lack of legislative rigor, but because it seems that the problems they address are beyond the ability of the legal system to solve; they will require more than just national legal solutions. In the end any solution to the surveillance problem will require international legal frameworks in addition to technological and cultural interventions.

Appendix: How It Works

In the appendix I have included (via link, because all four works are copyrighted and thus not able to be included within this book) other articles and videos that may be of interest to readers of this collection. *The Wall Street Journal's* series of articles on data mining does a thorough job of explaining the process by which corporations monitor internet traffic. The articles and accompanying video not only illustrate the use of third party cookies but demonstrate how one particular company uses them. A complementary video presentation by Ted Morgan explains how Skyhook location service tracks users' mobile devices. While in the video Morgan is mostly championing this technology as useful to consumers and corporations, it becomes pretty clear during his talk that not only is locative data becoming increasingly prevalent and powerful, but it also poses some serious privacy concerns.

Finally, I have linked to two full-length films available online. These two films explain the two sides of surveillance. In the first, *Erasing David*, filmmaker David Bond demonstrates the wealth of information that corporations and the government have collected about him. Bond's film invokes a 'man on the run' plot: he tries to avoid private investigators he has hired to attempt to find out everything about him. The ability of others to reconstruct his life with the data trails he leaves, as well as the expert interviews, are as enlightening as they are disturbing. Finally, the film *We Live in Public* appears to be about internet pioneer Josh Harris, but ultimately turns into a film about the interpersonal implications of living in a society where everyone is constantly surveilling everyone else. In the end, the film suggests that the most dangerous surveillance comes not from the government or corporations but from what we willingly accept as part of our social interactions.

References

Abelson, H. Ledeen, K. Lewis, H. (2008) *Blown to Bits*. Boston: Addison-Wesley .

Foucault, M. (1995) *Discipline and Punish: The Birth of the Prison.* New York:Vintage.

Vaidhyanathan, S. (2008) "Naked in the 'Nonopticon'" The Chronicle of Higher Education online, Feb. 15.

Articles

Knowing Everything: Data Mining

Somebody is Watching You: Video Surveillance

Monitoring Bodies: Surveilling Health

Daniel Chandramohan *et al.*
Should Data from Demographic Surveillance Systems Be Made More Widely Available to Researchers

Alireza Taravat Najafabad *et al.*
Web GIS and Public Health

Stuart Rennie *et al.*
Conducting Unlinked Anonymous HIV Surveillance in Developing Countries: Ethical, Epidemiological, and Public Health Concerns

Emily Chan *et al.*
Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance

Judging Privacy: Legal Issues

Samuel D. Warren, Louis D. Brandeis
The Right to Privacy

Daniel J. Solove
Data Mining and the Security-Liberty Debate

Omer Tene
What Google Knows: Privacy and Internet Search Engines

Paul Ohm
The Rise and Fall of Invasive ISP Surveillance

Appendix: How It Works

Emily Steel
A Web Pioneer Profiles Users by Name

The Wall Street Journal
Cracking the Code

Ted Morgan – Location Makes Mobile Mobile

David Bond
Erasing David

Ondi Timoner
We Live in Public

Attributions

AbuHmed, T. Mohaisen, A. Nyang, D. (2008). 'A Survey on Deep Packet Inspection for Intrusion Detection Systems,' *arXiv*. 0803.0037, 3, 2008. http://arxiv.org/abs/0803.0037v1.
Licence: © 2010 AbuHmed et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Artikis, A. Paliouras, G. (2009) 'A Logic Programming Approach to Behaviour Recognition,' *arXiv*. 0905.4614. 5, 2009. http://arxiv.org/abs/0905.4614v1.
Licence: © 2009 Artikis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Benabbas, Y., Ihaddadene, N. and Djeraba, C.(2011) 'Motion Pattern Extraction and Event Detection for Automatic Visual Surveillance,' *EURASIP Journal on Image and Video Processing*, Article ID 163682, 4. doi:10.1155/2011/163682.
Licence: © 2011 Benabbas et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Chan, E.H., Sahai, V., Conrad, C., Brownstein, J.S., (2011) 'Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance.' *PLoS Negl Trop Dis* 5(5): e1206. doi:10.1371/journal.pntd.0001206.
Licence: © 2011 Chan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Chandramohan D, Shibuya K, Setel P, Cairncross S, Lopez AD, et al. (2008) 'Should Data from Demographic Surveillance Systems Be Made More Widely Available to Researchers?' *PLoS Med* 5(2): e57. doi:10.1371/journal.pmed.0050057.
Licence: © 2008 Chandramohan. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

De Nardi, R. Holland, O. (2007). 'SwarMav: A Swarm of Miniature Aerial Vehicles.' Conference Presentation. *Cogprints*. 5569. 5, 2007. http://cogprints.org/5569/.
Licence: © 2007 De Nardi et al. Made available here via a link to the author's self-archived copy in the Cogprints repository.

Gayo-Avello, D. (2010). 'All Liaisons are Dangerous When all Your Friends are Known To Us.' *eprint arXiv.org*. 1012.5913, 12,
2010. http://arxiv.org/abs/1012.5913.

Licence: © 2010 Gayo-Avello. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Kholghi, M. Keyvanpour, M. (2011). 'An analytical framework for data stream mining techniques based on challenges and requirements.' *arXiv*. 1105.1950, 5, 2011. http://arxiv.org/abs/1105.1950.
Licence: © 2011 Khoglhi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Lee SH, Kim P-J, Ahn Y-Y, Jeong H, (2010). 'Googling Social Interactions: Web Search Engine Based Social Network Construction.' *PLoS ONE* 5(7): e11233. 3, 2010. doi:10.1371/journal.pone.0011233.
Licence: © 2010 Lee et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Najafabadi, AT. Pourhassan, M. (2010) 'Web GIS and Public Health Data.' *Online Journal of Health and Allied Sciences. Cogprints*. 6972. 9, 2010. http://cogprints.org/6972/.
Licence: © 2010 Najafabadi et. al. 'Made available here via a link to the author's self-archived copy in the Cogprints repository.

Ohm, Paul, (2008) 'The Rise and Fall of Invasive ISP Surveillance' (August 30, 2008). *University of Illinois Law Review*, 2009; U of Colorado Law Legal Studies Research Paper No. 08-22. Available at SSRN: http://ssrn.com/abstract=1261344.
Licence: © 2008 Ohm. Made available here via a link to the author's self-archived copy in the SSRN repository.

Pushpavathi, T.P., Selvarani, T.P., Shahsi, R. & Kumar, N.R. (2010). 'GPRS video Streaming Surveillance System GVSS,' *arXiv* 1002.3011. 2, 2010. http://arxiv.org/abs/1002.3011v1.
Licence: © 2010 Pushpavathi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Rennie, S., Turner, A.N., Mupenda, B., Behets, F., (2009) 'Conducting Unlinked Anonymous HIV Surveillance in Developing Countries: Ethical, Epidemiological, and Public Health Concerns.' *PLoS Med* 6(1): e1000004. doi:10.1371/journal.pmed.1000004
Licence: © 2009. Rennie et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Solove, D. J., (2008) 'Data Mining and the Security-Liberty Debate.' *University of Chicago Law Review* 74, p. 343; GWU Law School Public Law Research Paper No. 278. Available at SSRN: http://ssrn.com/abstract=990030.

Licence: © 2008 Solove. Made available here via a link to the author's self-archived copy in the SSRN repository.

Tene, O. (2007) 'What Google Knows: Privacy and Internet Search Engines,' Published online in draft form October 1; finally published in *Utah Law Review* 2008 (4). Available at SSRN: http://ssrn.com/abstract=1021490.
Licence: © 2007 Tene. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Warren, S. Brandeis, L. (1890). 'The Right to Privacy' *Harvard Law Review* IV (5). 12, 1890;[1].
License: In the public domain.

In the Appendix

Bond, D. (2009)*Erasing David*. Available online at: http://erasingdavid.com/

Morgan, T. (2010) 'Location Makes Mobile Mobile,' Momo Amsterdam Talk 1. Available online at: http://www.youtube.com/watch?v=bIJyWi9YsYU

Steel, E. (2010) "A Web Pioneer Profiles Users by Name." & 'Cracking the Code.' *Wall Street Journal*. October 25. Available online at: http://online.wsj.com/article/SB10001424052702304410504575560243259416072.html and http://s.wsj.net/public/resources/documents/st_RAPLEAF_20101018.html

Timoner, O. (2009) *We Live in Public*. Available online at: http://www.hulu.com/watch/192218/we-live-in-public