STATEMENT

of

Paul Rosenzweig
Red Branch Consulting, PLLC
Professorial Lecturer in Law, George Washington University
Visiting Fellow, The Heritage Foundation
Washington, D.C.

before the

Subcommittee on Oversight of Government Management, the Federal Workforce and the District of Columbia
Committee on Homeland Security and Governmental Affairs
United States Senate

July 31, 2012

**The State of Privacy and Security – Our Antique Privacy Rules**

**Introduction**

Chairman Akaka, Ranking Member Johnson, and Members of the Committee, I thank you for your invitation to appear today and present testimony on the question of data privacy and security under the Privacy Act. My name is Paul Rosenzweig and I am the Principal and founder of a small consulting company, Red Branch Consulting, PLLC, which specializes in, among other things, cybersecurity policy and legal advice. I am also a Senior Advisor to The Chertoff Group and a Professorial Lecturer in Law at George Washington University where I teach a course on Cybersecurity Law and Policy. In addition, I serve as a Visiting Fellow with a joint appointment in the Center for Legal & Judicial Studies and the Douglas and Sarah Allison Center for Foreign Policy Studies at The Heritage Foundation.[1] From 2005 to 2009 I served as the Deputy Assistant Secretary for Policy in the Department of Homeland Security.

[1] The Heritage Foundation is a public policy, research, and educational organization recognized as exempt under section 501(c)(3) of the Internal Revenue Code. It is privately supported and receives no funds from any government at any level, nor does it perform any government or other contract work.

The Heritage Foundation is the most broadly supported think tank in the United States. During 2011, it had nearly 700,000 individual, foundation, and corporate supporters representing every state in the U.S. Its 2011 income came from the following sources:

| | |
|---|---|
| Individuals | 78% |
| Foundations | 17% |

Needless to say, my testimony today is in my individual capacity and does not reflect the views of any institution with which I am affiliated or any of my various clients. Much of my testimony today is derived from prior academic work I have done in this field, most notably two research papers I published, one entitled "Privacy and Counter-Terrorism: The Pervasiveness of Data,"[2] and an older work entitled "Privacy and Consequences: Legal and Policy Structures for Implementing New Counter-Terrorism Technologies and Protecting Civil Liberty."[3] I used much of that research and additional work to create a Web-based project entitled "The Data Minefield"[4] while I was the Carnegie Visiting Fellow at the Medill School of Journalism, Northwestern University in 2011. All of that work, in turn, has been modified and will appear as part of several chapters in my forthcoming book, *Cyber Warfare: How Conflicts in Cyberspace are Challenging America and Changing the World* (Praeger Press 2012).

In my testimony today I want to make four basic points:

- The extent to which personal information is available due to society's increasing use and reliance on technology is growing every day. While one may view this as a good thing or a bad thing, it is, I submit, an inevitable thing. Wishing that it were not so is like King Canute commanding the tide not to come in. In the long run we do a disservice to our citizens if we do not recognize this reality.
- Thus, my second point is that it is, in my judgment, a mistake to speak of balancing privacy and information sharing in today's post-9/11 technological world. Rather, our objective should be to maximize both values. But this requires us to recognize that there is more than one way to protect privacy and that our current model of privacy is outdated and antiquated. Thus, while I am sure that all on this panel will agree that the Privacy Act needs to be updated, I suspect that my own views on how to do so are far more radical and transformative than those of my colleagues.
- In my view, the government can best ensure the privacy of the citizens by abandoning concepts like the Fair Information Practices that are tied to older technological conceptions. Instead of focusing on use and purpose limitations that are inconsistent with current capabilities and the threat environment (which requires the use of advanced data analytics) we would be better to focus privacy rules on the (admittedly more difficult) question of defining when it is and is not

---

Corporations 5%

The top five corporate givers provided The Heritage Foundation with 2% of its 2011 income. The Heritage Foundation's books are audited annually by the national accounting firm of McGladrey & Pullen. A list of major donors is available from The Heritage Foundation upon request.

Members of The Heritage Foundation staff testify as individuals discussing their own independent research. The views expressed are their own and do not reflect an institutional position for The Heritage Foundation or its board of trustees.

[2] 42 Case W. Res. J. Int'l L. 625 (2010).

[3] Robert Popp & John Yen, eds., *Emergent Information Technologies and Enabling Policies for Counter-Terrorism* (Wiley-IEEE 2006).

[4] http://nationalsecurityzone.org/datamining/.

appropriate to impose adverse consequences on citizens, combined with the equally essential (and also difficult) task of building a comprehensive oversight and audit system that constrains government activity effectively.

- It follows from what I've said already that I would not advise the Congress to undertake the task of updating the Privacy Act. Since I think that its entire structure is mismatched to technological reality, I would advocate a more extended consideration that leads to a complete rewrite of the statute along the lines I outline below.

**Dataveillance and Cyber Conflict**

Cyberspace is the natural battleground for enhanced analytical tools that are enabled by the technology of data collection. If our goal is to combat terrorists or insurgents (or even other nations) then the cyber domain offers us the capacity not just to steal secret information through espionage, but to take observable public behavior and information and use cyber tools to develop a more nuanced and robust understanding of their tactics and intentions. Likewise, it can be used by our opponents to uncover our own secrets.

Traditionally, the concept of "surveillance" has been taken to mean an act of physical surveillance—e.g., following someone around or planting a secret camera in an apartment. As technology improved, our spy agencies and law enforcement institutions increasingly came to rely on even more sophisticated technical means of surveillance,[5] and so we came to develop the capacity to electronically intercept telecommunications and examine email while in transit.[6]

To these more "traditional" forms of surveillance we must now add another: the collection and analysis of personal data and information about an individual or organization. Call the phenomenon "dataveillance" if you wish, but it is an inevitable product of our increasing reliance on the Internet and global communications systems. One leaves an electronic trail almost everywhere you go. Increasingly, in a networked world technological changes have made personal information pervasively available. As the available storehouse of data has grown, so have governmental and commercial efforts to use this personal data for their own purposes. Commercial enterprises target ads and solicit new customers. Governments use the data to, for example, identify and target previously unknown terror suspects—to find so-called clean skins who are not in any intelligence database. This capability for enhanced data analysis has already proven its utility and holds great promise for the future of commercial activity and counter-terrorism efforts.

[5] For an overarching history of the transition from human intelligence to U-2 spy planes and, eventually, to satellites, *see generally* Tim Weiner*, Legacy of Ashes: The History of the CIA* (2007).

[6] Law enforcement electronic interceptions are generally governed by Title III of the Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, 82 Stat. 197 (codified in scattered sections of 5, 18, and 42 U.S.C.), and intelligence interceptions are governed by the Foreign Intelligence Surveillance Act of 1978, Pub. L. No. 95-511, 92 Stat. 1783 (codified as amended in scattered sections of 50 U.S.C.).

Yet this analytical capacity also comes at a price—the peril of creating an ineradicable trove of information about innocent individuals. That peril is typically supposed to stem from problems of misuse; in the government sphere one imagines data mining to identify political opponents, and in the private sector we fear targeted spam. To be sure, that is a danger to be guarded against.

But the dangers of pervasively available data also arise from other factors. Often, for example, there is an absence of context to the data that permits or requires inaccurate inferences. Knowing that an individual has a criminal conviction is a bare data point; knowing what the conviction was for and in what context allows for a more granular and refined judgment.

The challenges arising from these new forms of analysis have already become the subject of significant political debate. One need but think of the controversy surrounding the most ambitious of these—the Total Information Awareness (TIA) program. TIA was a research program initiated by the Defense Advanced Research Projects Agency (DARPA) in the immediate aftermath of September 11. Its conception was to use advanced data analysis techniques to search the information space of commercial and public sector data looking for threat signatures that were indicative of a terrorist threat. Because it would have given the government access to vast quantities of data about individuals, it was condemned as a return of "Big Brother."[7]

Compare that condemnation with the universal criticism of the government for its failure to "connect the dots" during the Christmas 2009 bomb plot attempted by Umar Farouk Abdulmutallab.[8] This gives you some idea of the crosscurrents at play. The conundrum arises because the analytical techniques are fundamentally similar to those used by traditional law enforcement agencies, but they operate on so much vaster a set of data, and that data is so much more readily capable of analysis and manipulation, that the differences in degree tend to become differences in kind. To put the issue in perspective, just consider a partial listing of relevant databases that might be targeted: credit card, telephone calls, criminal records, real estate purchases, travel itineraries, and so on.

One thing is certain—these analytical tools are of such great utility that governments will expand their use, as will the private sector. Old rules about collection and use limitations are no longer technologically relevant. If we value privacy at all, these ineffective protections must be replaced with new constructs. The goal then is the identification of a suitable legal and policy regime to regulate and manage the use of mass quantities of personal data.

---

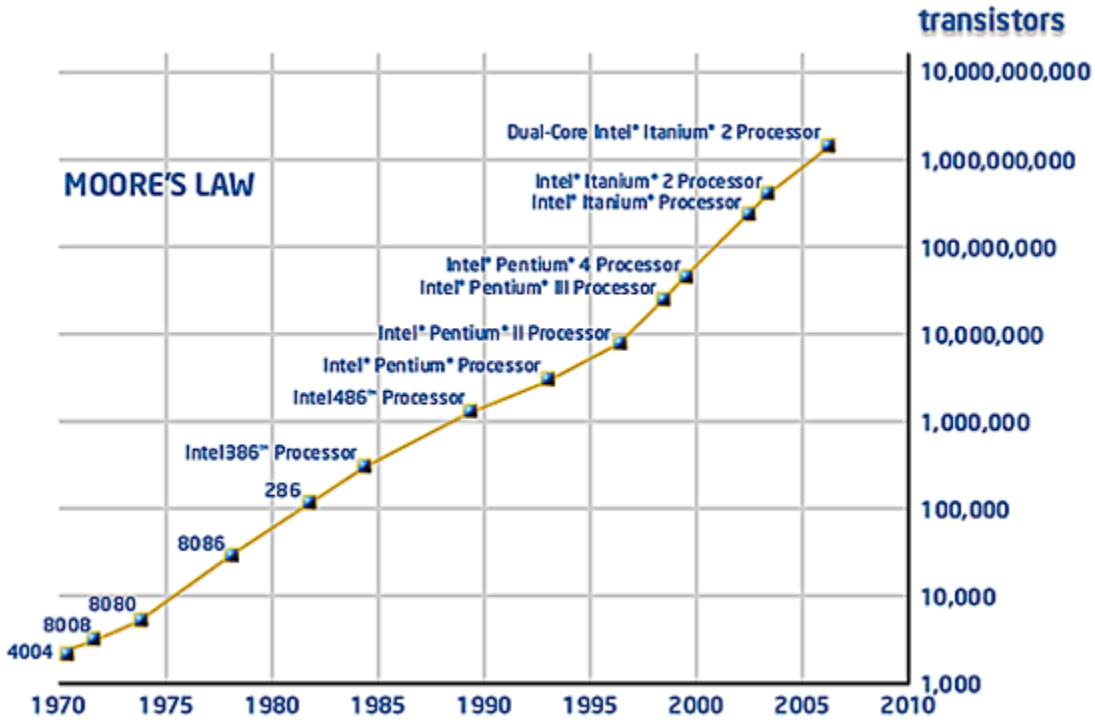
[7] An article by William Safire instigated a significant political controversy. *See* William Safire, "You Are a Suspect," *The New York Times*, Nov. 14, 2002, at A35. It led directly to the creation of a blue-ribbon panel, the Technology and Privacy Advisory Committee, and, eventually, to the cancellation of the Total Information Awareness program. The final report of the Technology and Privacy Advisory Committee is available at http://www.defense.gov/news/Jan2006/d20060208tapac.pdf (last visited Feb. 23, 2010).

[8] *See*, *e.g.*, Scott Shane & Eric Lipton, "Passengers' Actions Thwart a Plan to Down a Jet," *The New York Times,* Dec. 27, 2009, at A1.

**The Computing and Storage Revolution**

The growth of dataveillance is inevitable. It reflects a fundamental change caused by technological advances that, like King Canute's fabled tide, cannot be stopped or slowed. Increasingly, the cyber conflict will be fought, and won, by those who use data to their best advantage. The opportunity—or problem, depending on one's perspective—derives from two related, yet distinct trends: increases in computing power and decreases in data storage costs.

Many are familiar with the long-term increase in the power of computers. It is most familiarly characterized as Moore's Law—named after Intel computer scientist Gordon Moore, who first posited the law in 1965. Moore's Law predicts that computer chip capacities will double every eighteen to twenty-four months.[9] Moore's law has been remarkably constant for nearly thirty years, as the graph below demonstrates.[10]

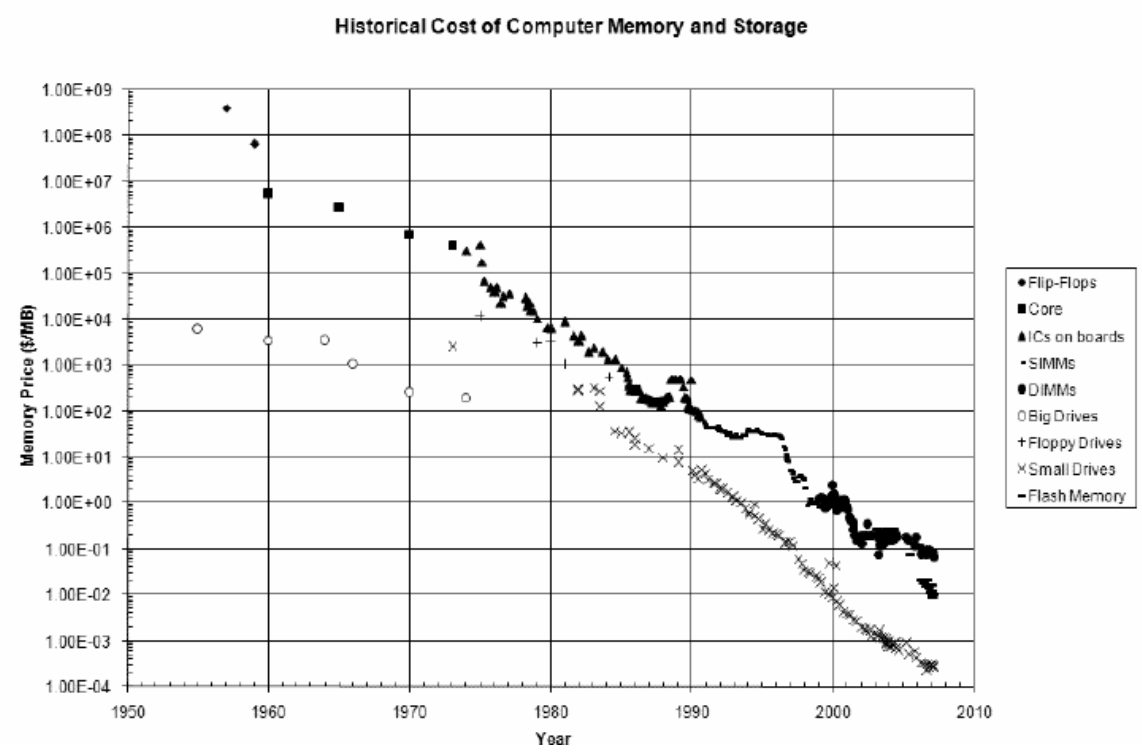



The scale makes clear that the effect of routine doubling is logarithmic. Processor capacity today is roughly more than one million times faster than processor speed in 1970.

[9] *See* Linda Null & Julia Lobur, *The Essentials of Computer Organization and Architecture* 27 (2d ed. 2006).
[10] Charts of Moore's law are widely available. This one is from http://www.deepspar.com/images/MooresLaw.jpg (last visited Feb. 23, 2010).

The power of this processing capacity—which translates almost directly into processing speed—is immense. And though no one predicts that processing speed will double indefinitely—surely a physical impossibility—there is no current expectation that the limits of chip capacity have been reached.

To this trend one must also add the remarkable reduction in the costs of data storage. As the following chart demonstrates,[11] data storage costs have also been decreasing at a logarithmic rate, almost identical to the increases we have experienced in chip capacity, but with an inverse slope.




What this means in practical terms is that in 1984— less than thirty years ago—it cost roughly two hundred dollars to store a megabyte of data. By 1999 that cost had sunk to seventy-five cents. Today you can buy one hundred megabytes of data storage capacity for a penny. On eBay you can frequently purchase a terabyte storage device for your desktop for under one hundred dollars. A terabyte is roughly 1 trillion bytes of data—a huge volume for storing simple alphanumeric information. Here, too, the prospects are for ever-cheaper data storage. One can readily imagine peta-, exa-, or even yottabyte sized personal storage devices.[12] If that is for the individual, imagine what a large corporation or a government can purchase and maintain.

[11] Lev Lafayette, "Definition, History, Usage and Future of Computer Data Storage," *Organdi,* http://organdi.net/article.php3?id_article=82 (the graph is directly available at http://organdi.net/IMG/gif/historical_cost_graph5.gif).

[12] A petabyte is $1000^5$ bytes, a exabyte is $1000^6$ bytes, and a yottabyte is $1000^8$ bytes.

Therefore, the story of technology today requires us to answer the question: "What happens when ever-quicker processing power meets ever-cheaper storage capacity?" Anyone who uses Gmail knows the answer to that question. No longer do you have to laboriously label, file, and tag your email. One may now simply store all the email he or she wants to retain and use a simple natural language search algorithm to pull up relevant emails from storage when needed. The storage cost of Gmail to the user is zero—Google offers it for free—and the processing time for any search request for the average individual is measured in, at most, seconds, not minutes.

Here is how IBM Chairman Samuel J. Palmisano put it in a speech he gave in September 2011:

> We're all aware of the approximately two billion people now on the Internet—in every part of the planet, thanks to the explosion of mobile technology.
>
> But there are also upwards of a trillion interconnected and intelligent objects and organisms—what some call the Internet of Things.
>
> All of this is generating vast stores of information. It is estimated that there will be 44 times as much data and content coming over the next decade...reaching 35 zettabytes in 2020. A zettabyte is a 1 followed by 21 zeros. And thanks to advanced computation and analytics, we can now make sense of that data in something like real time. This enables very different kinds of insight, foresight and decision-making.[13]

In other words, we live in the world of "Big Data." Data is now pervasively available and pervasively searchable. For large-scale databases of the size maintained by governments or companies, the practical limitations lie in the actual search algorithms used and how they are designed to process the data, not in the chips or the storage units. The changes that will come from this new cyber reality are profound.

**The Power of Data Analytics**

Ten years ago, surveying the technology of the time—which, by and large, was one hundred times *less* powerful than today's data processing capacity—Scott McNealy, then-CEO of Sun Microsystems, said, "Privacy is dead. Get over it."[14] He was, it seems, slightly wrong. Pure privacy—that is, the privacy of activities in your own home—remains reasonably well-protected.[15] What has been lost, and will become even more so increasingly, is the anonymity of being able to act in public (whether physically or in cyberspace) without anyone having the technological capacity to permanently record and retain data

[13] Samuel J. Palmisano, "Thoughts on the Future of Leadership," September 20, 2011, https://www.ibm.com/smarterplanet/us/en/leadership/stories/pdf/prepared_remarks.pdf.
[14] Though the original statement may be apocryphal, many have quoted it since, including McNealy himself. *See*, *e.g.*, Matt Hamblen, "McNealy Calls for Smart Cards," *Computer World*, Oct 12, 2001, http://www.computerworld.com/s/article/64729/McNealy_calls_for_smart_cards_to_help_security.
[15] *See*, *e.g.*, Kyllo v. United States, 533 U.S. 27 (2001) (the use of thermal imagining outside the home without a warrant is an illegal search when it is used, even indirectly, to reveal activity taking place within the home).

about your activity for later analysis. Today, large data collection and aggregation companies, such as Experian and Axicom, may hire retirees to harvest, by hand, public records from government databases.[16] Paper records are digitized and electronic records are downloaded. These data aggregation companies typically hold birth records, credit and conviction records, real estate transactions and liens, bridal registries, and even kennel club records. One company, Acxiom, estimates that it holds on average approximately 1,500 pieces of data on each adult American.[17]

Since most, though not all, of these records are governmental in origin, the government has equivalent access to the data, and what they cannot create themselves they can likely buy or demand from the private sector. The day is now here when anyone with enough data and sufficient computing power can develop a detailed picture of any identifiable individual. That picture might tell your food preferences or your underwear size. It might tell something about your terrorist activity. Or your politics.

This analytical capacity can have a powerful influence in law and policy—and in particular in revealing links between the cyber personas and the real world activities of individuals. When we speak of the new form of "dataveillance," we are not speaking of the comparatively simple matching algorithms that cross check when a person's name is submitted for review—when, for example, they apply for a job. Even that exercise is a challenge for any government, as the failure to list Abdulmutallab in advance of the 2009 Christmas bombing attempt demonstrates.[18] The process contains uncertainties of data accuracy and fidelity, analysis and registration, transmission and propagation, and review, correction, and revision. Yet, even with those complexities, the process uses relatively simple technologically—the implementation is what poses a challenge.

By contrast, other systems of data analysis are far more technologically sophisticated. They are, in the end, an attempt to sift through large quantities of personal information to identify subjects when their identities are not already known. In the commercial context, these individuals are called "potential customers." In the cyber conflict context, they might be called "Anonymous" or "Russian patriotic hackers." In the terrorism context, they are often called "clean skins" because there is no known derogatory information connected to their names or identities. In this latter context, the individuals are dangerous because nothing is known of their predilections. For precisely this reason, this form of data analysis is sometimes called "knowledge discovery," as the intention is to discover something previously unknown about an individual. There can be little doubt that data analysis of this sort can prove to be of great value. A few examples will illustrate the point.

[16] I learned this from discussions with ChoicePoint's CEO Derek Smith and other industry practitioners. *See also* Ralph M. Stair & George W. Reynolds, *Fundamentals of Information Systems* 362 (2003) (discussing Experian's collection of public records from government databases).

[17] Stephanie Clifford, "Online Ads Follow Web Users, and Get Much More Personal," *The New York Times,* July 30, 2009, at A1.

[18] Peter Baker & Carl Hulse, "Obama Hears of Signs That Should Have Grounded Plot," *The New York Times,* Dec. 30, 2009, at A1.

The story of Ra'ed al-Banna, a Jordanian who attempted to enter the U.S. at O'Hare Airport on June 14, 2003, illustrates the value of computer dataveillance.[19] al-Banna was carrying a valid business visa in his Jordanian passport and, on the surface, appeared to be an unremarkable business traveler from the Middle East.

The Department of Homeland Security operates a sophisticated data analysis program called the Automated Targeting System (ATS) to assess the comparative risks of arriving passengers. Based on those assessments, the inspection resources of Customs and Border Protection (CBP) are allocated.[20] The system is essential given the sheer volume of travelers to America. In a typical year approximately three hundred and fifty million people sought entry across our borders, and more than eighty-five million of those arrived by air.[21] Since over three hundred and fifty million individuals cannot, obviously, be subject to intense scrutiny, some form of assessment and analysis must be used to make choices about how and when to conduct inspections. ATS is that system.

ATS flagged al-Banna for heightened scrutiny.[22] His pattern of travel and his prior record of entry to the U.S. combined to suggest that he should be subjected to secondary screening—a form of enhanced, individualized review where a passenger is pulled from the main line of entrants and individually questioned. During the secondary interview, al-Banna's answers were inconsistent and evasive—so much so that the CBP officer who conducted the interview decided to deny his application for entry and ordered him returned to his point of origin.[23] As a matter of routine, al-Banna's photograph and fingerprints were collected before he was send on his way.

There the story might have ended, since CBP officers reject entry applications daily for a host of reasons, but al-Banna proved an unusual case. More than a year later, in February 2005, a car filled with explosives drove into a crowd of military and police recruits in the town of Hillah, Iraq.[24] More than one hundred twenty-five people died—the largest death toll for a single incident in Iraq until that time. The suicide bomber's hand and forearm were found chained to the steering wheel of the exploded car (why they were chained is a fascinating question of psychology). When the fingerprints were taken by U.S.

---

[19] A summary of the al-Banna case can be found in Stewart A. Baker & Nathan A. Sales, "Homeland Security, Information Policy, and the Transatlantic Alliance," *in* George Mason University Law and Economics Research Paper Series 09-20 (March 2009), http://ssrn.com/abstract=1361943. *See also* Charlotte Buchen, *The Man Turned Away*, PBS FRONTLINE, Oct. 10, 2006, www.pbs.org/wgbh/pages/frontline/enemywithinh/reality/al-banna.html.

[20] For a more thorough description of the ATS, *see* Paul Rosenzweig, "Targeting Terrorists: The Counterrevolution," 34 Wm. Mitchell L. Rev. 5083, 5086–90 (2008). *See also* Privacy Act of 1974, Notice of Privacy Act System of Records, 72 Fed. Reg. 43,650–02 (Aug. 6, 2007) (providing details of the ATS).

[21] *See* Customs and Border Protection, On a Typical Day in Fiscal Year 2009, CBP . . ., http://www.cbp.gov/xp/cgov/about/accomplish/fy09_typical_day.xml.

[22] *See* Scott Shane & Lowell Bergman, "Contained? Adding Up the Ounces of Prevention," *The New York Times,* Sep. 10, 2006, § 4, at 1.

[23] U.S. Customs and Border Protection, CBP: Securing America's Borders 4 (Sept. 2006), http://www.customs.gov/linkhandler/cgov/newsroom/publications/mission/cbp_securing_borders.ctt/cbp_securing_borders.pdf.

[24] *See* Shane & Bergman*, supra.*

military forces, a match was found to the fingerprints taken from al-Banna twenty months earlier in Chicago.

Now, of course, nobody knows what al-Banna intended to do that day when he arrived at O'Hare. It is impossible to prove a counterfactual. Perhaps he was only headed to visit friends, but the CBP officer who interviewed al-Banna later said, "I was shocked. That it was so close to home, that I actually interviewed someone who not only was capable of doing but actually did something like that. You never know who you are interviewing or what they are capable of doing."[25] Without the data analysis provided by ATS, it is nearly certain that al-Banna would have entered the U.S.—who knows for what purpose.

Most similar successes are not made public. Often the factors that form part of the analysis cannot be revealed, and successes in identifying terrorist suspects—or, in other contexts, members of a criminal organization—would be negated by disclosure of the success. Only al-Banna's death made his case fit for public disclosure.

That does not mean that a careful observer cannot discern the outlines of other cyber intelligence successes based on data analysis in recent events. When David Headley was arrested for allegedly seeking to commit terrorist acts in Denmark, news reports suggested that one of the key factors in his identification was his pattern of travel to the Middle East and his efforts to conceal those trips from the government.[26] Dataveillance of his travel provided both the trigger to ask questions and the factual cross-check on the veracity of his answers. Likewise, when Najibullah Zazi (who tried to explode a bomb in Times Square) was arrested, one factor that was publicly disclosed as a ground for suspicion was his travel to Pakistan.[27]

Both of these incidents, which involved serious threats of violence, would appear to have been thwarted, at least in part, through some form of successful dataveillance, i.e., using knowledge discovery techniques to target investigative resources based upon a careful risk assessment of seemingly innocent individuated facts.

Our failures also seem to arise when these sorts of cyber analytic techniques are used ineffectively. In the case of the 2009 Christmas bomb plot, not only was Abdulmutallab's name provided by his father, but the evidence suggests that other, less specific NSA intercepts existed that might have generated a suspicion of Nigerian travelers.[28] Add in his reported purchase of a ticket with cash and the alleged

[25] DHS Success Stories Case # 000016 (2005/03/01) (on file with author).
[26] *See* Cam Simpson & Siobhan Gorman, "Terror Suspect Failed a Test," *Wall St. Journal*, Dec. 9, 2009, at A4.
[27] For example, the Department of Justice's Motion for a Permanent Order of Detention cites CBP records of trips to Pakistan. Memorandum of Law in Support of the Government's Motion for a Permanent Order of Detention at 3–4, United States v. Najibullah Zazi, No. 09-CR-663 (RJD) (E.D.N.Y. Sept. 24, 2009), http://www.justice.gov/opa/documents/zazi-detention-memo.pdf.
[28] Umar Farouk Abdulmutallab, http://topics.nytimes.com/top/reference/timestopics/people/a/umar_farouk_abdulmutallab/index.html.

rejection of his visa application by the U.K.[29] and the case seems to be the precise sort of concatenation of facts which, individually, amount to little but, collectively, paint a more cautionary picture. In the wake of the failed bombing attempt, there are already calls for even greater efforts to "connect the dots" of terrorist threats and that will mean more dataveillance, not less.[30]

**Antique Privacy**

Cyber dataveillance is here to stay whether we like it or not. The only question is when and how we monitor and control the government's use of the techniques so that we get the benefits of the growth in data surveillance without the potential harms to civil liberties.

As should be evident, the use of such analytical tools is not without risks. The same systems that sift layers of data to identify concealed terrorist links are just as capable, if set to the task, of stripping anonymity from many other forms of conduct—personal purchases, politics, and peccadilloes. The question then becomes how do we empower data analysis for good purposes while providing oversight mechanisms for deterring malfeasant uses?

Our current privacy-protective architecture, or, if one prefers, our anonymity-protective architecture, is simply not up to the task. It is, to a very real degree, an antique relic of the last century. The relevant Supreme Court precedents date from the 1970s, as does the 1974 Privacy Act.[31] Is it any wonder that the current structure of law does not match the technological reality?

The "third party doctrine" developed by the Supreme Court in two 1970-era cases—*United States v. Miller*[32] and *Smith v. Maryland*[33]—at the dawn of the computer era, means that information you disclose to a third party is not protected by the Fourth Amendment. In the context of data privacy, that means that there is no constitutional protection against the collection and aggregation of your cyber data (credit card purchase and the like) for purposes of data analysis and piercing the veil of anonymity.[34]

---

[29] *Id.*; John F. Burns, "Britain Says Bomb Suspect Was Denied Visa Renewal," *The New York Times,* Dec. 29, 2009, at A12.

[30] *See* Ben Feller, "Obama: The Buck Stops with Me," *Huffington Post*, Jan. 7, 2010, http://www.huffingtonpost.com/2010/01/07/obama-christmas-bomber-report_n_414309.html.

[31] 5 U.S.C. § 552a (2006).

[32] 425 U.S. 435 (1976).

[33] 442 U.S. 735 (1979).

[34] I should note here an important qualification. In January 2012, the Supreme Court decided *United States v. Jones*, __ U.S. ___ (No. 10-1259, Jan. 23, 2012), http://www.supremecourt.gov/opinions/11pdf/10-1259.pdf. The two concurring opinions in the case suggest that at some future point the Court may revisit the third-party doctrine. For now, however, as Congress considers its legislative options, the existing legal architecture remains unchanged and I take it as settled … until, of course, it no longer is.

At the federal level, what protects anonymity are the statutory protections created by Congress.[35] Some laws, like the Right to Financial Privacy Act (RFPA),[36] create sector-specific privacy protections. Reacting to *Miller*, the RFPA prevents banks from willy-nilly providing financial data to the government, instead requiring the issuance of a subpoena and notice to a customer who has the right to object to the inquiry. Likewise, the Health Insurance Portability and Accountability Act[37] has stringent rules regarding medical privacy and limiting the types of disclosures that doctors, hospitals, and insurers can make.

By and large, however, in the national security dataveillance sphere there is no sector or activity-specific set of protections.[38] Rather, we seek to protect privacy (or anonymity) by requiring the government to adhere to broad principles of privacy protection. These principles, known as the Fair Information Principles,[39] were first developed in the U.S. and have now become the touchstone of most privacy protective regimes. They are embedded in the Privacy Act of 1974 and lie at the core of the European Union's 1995 Privacy Directive.[40] In brief summary—which does not do them justice for want of detail—the principles are:

- *Collection limitation*: The collection of personal information should be lawful and limited to that which is necessary. Where feasible, the collection should be consensual.
- *Data quality*: Those collecting information should strive to ensure that it is accurate, relevant, and complete.
- *Purpose specification*: Data should be collected for a specific purpose. Data should not be repurposed to other uses without disclosure and consent, if at all.
- *Use limitation*: Data should be used only for a specific purpose and should be disclosed only for the purpose collected.
- *Security safeguards*: Information collected should be protected against loss or theft.
- *Openness*: The collection, use, and security of data collected should be fully disclosed and transparent to the public.
- *Individual participation*: Individuals should be allowed to access data collected about themselves and should be afforded a chance to correct any errors they perceive.
- *Accountability*: Those who collect and hold data should be accountable for their adherence to these norms.[41]

[35] There exist state-based statutory privacy protections and most state courts recognize a common law right to privacy of some form. *See* Samuel Warren & Louis D. Brandeis, "The Right to Privacy," 4 Harv. L. Rev. 193 (1890). Neither is an effective limitation on the action of the federal government.
[36] 12 U.S.C. §§ 3401–3422 (2006).
[37] Pub. L. No. 104-191, 110 Stat. 1936 (codified in scattered sections of 26, 29, and 42 U.S.C.).
[38] The Foreign Intelligence Surveillance Act is a notable exception, governing the collection of the substance (as opposed to the call record data) of personal communications. *See* Foreign Intelligence Surveillance Act, 50 U.S.C. §§ 1801–1871 (2006).
[39] *See* Privacy Rights Clearinghouse, "A Review of the Fair Information Principles: The Foundation of Privacy Public Policy," http://www.privacyrights.org/ar/fairinfo.htm.
[40] *See* Privacy Act, 5 U.S.C. § 552a (2006); Council Directive 95/46/EC, 1995 O.J. (L281) 31, http://ec.europa.eu/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf.
[41] *See* Fair Information Principles, *supra*.

In the U.S., these principles are procedurally implemented through Privacy Impact Assessments (PIAs) and through the publication of System of Record Notices (SORNs).[42] The PIA, conducted by the government, is a detailed analysis of how a particular set of personal information is collected, stored, protected, shared, and managed. The SORN is the public notification of the existence of systems that collect and hold data. Taken together, the two requirements are intended to provide for the openness and accountability that will allow the public to remain assured that those collecting data are adhering to these principles.[43]

The problem is that a conscientious and fair application of these principles is, in many ways, fundamentally inconsistent with the way in which personal information can be used in the context of counter-terrorism or cyber insurgency dataveillance. Recognizing this fact is not, at this juncture, to make a normative judgment, but merely to make the descriptive point that the way in which dataveillance programs, like the Automated Targeting System that discovered al-Banna, function is at odds with these principles.

Consider that the collection limitation principle calls for the collection of the least amount of information and, where feasible, acquiring the consent of those about whom the data is being collected. Effective terrorism dataveillance, however, relies on the breadth of the collection for its success since the unknown connection will often come from an unexpected data field and the collection often occurs without the knowledge of, much less the consent of, the data subject.

Likewise, the purpose specification principle, if fully applied, would significantly degrade the analytical utility of many knowledge discovery systems. Often the data of interest that gives rise to a previously unknown connection is one that was collected for a different purpose and intended for a different use. To take the most prosaic example, imagine that a phone number is collected from an air traveler so that the airline may contact him, and his frequent flyer number is collected so that his loyalty account may be credited. When those data fields are used for another purpose—for example, to identify potential connections between known terrorists and those who are otherwise unknown—these purpose and use limitation principles are violated. Yet that is precisely how systems like ATS operate and, in retrospect, it is a method that might have identified the 9/11 terrorists before their attack if it had been available at the time.[44]

---

[42] *See* U.S. Securities and Exchange Commission, *Privacy Impact Assessment (PIF) Guide* 4 (Jan. 2007), www.sec.gov/about/privacy/piaguide.pdf.

[43] Separately, the Privacy Act also affords individuals the right to go to court to correct erroneous data collected about them. 5 U.S.C. § 552a(d) (2006). It is a never-ending source of friction with our international partners that this right extends only to American citizens and legal residents.

[44] *See* Newton N. Minow, "Seven Clicks Away," *Wall St. Journal,* , June 3, 2004, at A14; The Markle Foundation, *Protecting America's Freedom in the Information Age: A Report of the Markle Foundation Task Force* 28 (2002), http://www.markle.org/downloadable_assets/nstf_full.pdf.

Perhaps even more pointedly, the principles of openness and individual participation are challenging to implement in the counter-terror context. Full disclosure of the methods of operation of a dataveillance system would often make it easier, for those wishing to do so, to evade it. The notion of allowing potential terrorists to see exactly what data is and is not held about them simply seems impossible to contemplate.

The problem, of course, is that in this modern world of widely distributed networks with massive data storage capacity and computational capacity, so much analysis becomes possible that the old principles no longer fit. We could, of course, apply them, but only at the cost of completely disabling the new analytic capacity. In the current time of cyber threat that seems unlikely. Alternatively, we can abandon privacy altogether, allowing technology to run rampant with no control. That, too, seems unlikely and unwise.

What is needed, then, is a modernized conception of privacy—one with the flexibility to allow effective government action but with the surety necessary to protect against government abuse.

**Modernizing Privacy**

Our privacy laws and our conceptions of privacy cannot withstand the technological change that is happening and the cyber conflict that is developing. We must put theories of data availability and anonymity on a sounder footing—a footing that will withstand the rigors of ever-increasing computational capacity. To do so we need to define what values underlie our instinctive privacy-protective reaction to the new technology, assess how realistic threats of abuse and misuse are, and create legal and policy incentives to foster positive applications while restraining adverse ones.

Though a comprehensive new anonymity-protective legal structure has yet to be developed, the outline of one can already be discerned. Old ideas of collection and purpose limitations will be forced by technological change to yield to a greater emphasis on use limitations. Even those limitations will need to be modified so that our concern is not with uses that are mere "analyses" but rather with uses that constitute the "imposition of adverse consequences." The new system will be based on the new answers to three broad questions:

- What is privacy?
- What new structural systems do we need?
- What old rules need to be rethought?

*What is Privacy?* —Privacy is really a misnomer. What it reflects is a desire for independence of personal activity, a form of autonomy. We protect that privacy in many ways. Sometimes we do so through secrecy which effectively obscures both observation of conduct and the identity of those engaging in the conduct. In other instances we protect the autonomy directly. Even though conduct is observed and the

actor identified, we provide direct rules to limit action—as, for example, in the criminal context where we have an exclusionary rule to limit the use of illegally collected evidence.

The concept of privacy that most applies to the new information technology regime is the idea of anonymity or "practical obscurity," a middle ground where observation is permitted—that is, we expose our actions in public—but we are not subject to identification or scrutiny. The information data-space is suffused with information of this middle-ground sort, e.g., bank account transactions, phone records, airplane reservations, and Smartcard travel logs to name but a few. They constitute the core of transactions and electronic signature or verification information available in cyberspace. The anonymity that one has in respect of these transactions is not terribly different from "real-world anonymity." Consider, as an example, the act of driving a car. It is done in public, but one is generally not subject to routine identification and scrutiny.

Protecting the anonymity we value requires, in the first instance, defining it accurately. One might posit that anonymity is, in effect, the ability to walk through the world unexamined. That is, however, not strictly accurate, for our conduct is examined numerous times every day. Sometimes the examination is by a private individual—for example, one may notice that the individual sitting next to them on the train is wearing a wedding ring. Other routine examinations are by governmental authorities—the policeman in the car who watches the street or the security camera at the bank or airport, for example. As we drive down the road, any number of people might observe us.

So what we really must mean by anonymity is not a pure form of privacy akin to secrecy. Rather, what we mean is that even though one's conduct is examined, routinely and regularly, both with and without one's knowledge, *nothing adverse should happen to you without good cause*. In other words, the veil of anonymity—previously protected by our "practical obscurity"—that is now so readily pierced by technology must be protected by rules that limit when the piercing may happen as a means of protecting privacy and preventing governmental abuse. To put it more precisely, the key to this conception of privacy is that privacy's principal virtue is a *limitation on consequence*. If there are no unjustified consequences—i.e., consequences that are the product of abuse or error or the application of an unwise policy—then, under this vision, there is no effect on a cognizable liberty/privacy interest. In other words, if nobody is there to hear the tree, or identify the actor, it really does not make a sound.

The appeal of this model is that it is, by and large, the model we already have for government/personal interactions in the physical world. The rule is not that the police cannot observe you; it is that they require authorization of some form from some authority in order to be permitted to engage in certain types of interactions, which are identified here as "consequences." The police normally cannot stop you to question you without "reasonable suspicion," cannot arrest you without "probable cause," cannot search your house without "probable cause," and cannot examine a corporation's business records about you without a showing of "relevance" to an ongoing investigation. We can and should build structures that map the same rules-based model of authorization linked to consequence as the appropriate model for the world of dataveillance.

Thus, the questions to be asked of any dataveillance program are: What is the consequence of identification? What is the trigger for that consequence? Who decides when the trigger is met? These questions are the ones that really matter, and questions of collection limitation or purpose limitation, for example, are rightly seen as distractions from the main point. The right answers to these questions will vary, of course, depending on the context of the inquiry, but the critical first step is making sure that we are asking the right questions.

*What New Structural Systems Do We Need?* —Once defined, how do we protect anonymity?[45] The traditional way is with a system of rules and a system of oversight for compliance with those rules. Here, too, modifications need to be made in light of technological change.

Rules, for example, tend to be static and unchanging and do not account readily for changes in technology. Indeed, the Privacy Act—the central statute intended to protect individual privacy against government intrusion—is emblematic of this problem; the principles of the Privacy Act are ill-suited to most of the new technological methodologies, such as distributed databases. Thus, we have begun to develop new systems and structures.

First, we are changing from a top-down process of command and control rule to one in which the principal means of privacy protection is through institutional oversight. To that end, the Department of Homeland Security was created with a statutorily required Privacy Officer (and another Officer for Civil Rights and Civil Liberties).[46] The more recent Intelligence Reform and Terrorism Prevention Act,[47] and the Implementing Recommendations of the 9/11 Commission Act of 2007[48] go further. For the first time, they created a Civil Liberties Protection Officer within the intelligence community. More generally, intelligence activities are to be overseen by an independent Privacy and Civil Liberties Oversight Board.[49] Indeed, these institutions serve a novel dual function. They are, in effect, internal watchdogs for privacy concerns. In addition, they naturally serve as a focus for external complaints, requiring them to exercise some of the function of ombudsmen. In either capacity, they are a new structural invention on the American scene—at least, with respect to privacy concerns.

Second, and perhaps most significantly, the very same dataveillance systems that are used to advance our counter-terrorism interests are equally well suited to assure that government officials comply with

[45] This section is based in part on the essay Paul Rosenzweig, "The Changing Face of Privacy Policy and the New Policy-Technology Interface," *IEEE Intelligent Systems, Trends and Controversies* 84–86 (Sept.–Oct. 2005), www.dartmouth.edu/~humanterrain/papers/intelligent_systems.pdf.
[46] *See* Homeland Security Act of 2002, Pub. L. No. 107-296 § 222 (2002).
[47] *See* Intelligence Reform and Terrorism Prevention Act of 2004, Pub. L. No. 108-458, 118 Stat. 3638.
[48] Implementing Recommendations of the 9/11 Commission Act of 2007, Pub. L. No. 110-53, § 1502, 121 Stat. 266, 424 (codified at 6 U.S.C.A. § 1152(g) (West 2008)).
[49] The duties of Civil Liberties and Privacy Officer in the Office of the Director of National Intelligence are codified at 50 U.S.C. § 403–3d (2006). The Privacy and Civil Liberties Oversight Board is authorized by section 801 of the Implementing Recommendations of the 9/11 Commission Act of 2007.

the limitations imposed on them in respect of individual privacy. Put another way, the dataveillance systems are uniquely well equipped to watch the watchers, and the first people who should lose their privacy are the officials who might wrongfully invade the privacy of others.

Indeed, there are already indications that these strong audit mechanisms are effective. Recall the incident in the last presidential campaign in which contractors hacked Barack Obama's passport file.[50] In this instance, there was no lawful reason for the disclosure of the file; it was disclosed purely for prurient, political reasons. As a result, candidate Obama suffered an adverse consequence of disclosure which had not met any legal trigger that would have permitted the disclosure. A strong audit function quickly identified the wrongdoers and allowed punitive action to be taken.[51]

We can, therefore, be reasonably confident that as we move forward in establishing a consequence-based system of privacy protection we are also moving toward a point where the legal structures and technological capabilities to support that system are being put into place.

*What Old Rules Need to Be Rethought?* —Perhaps the greatest dangers, however, lie in questions that we have yet to ask—at least those that have not yet been heard.[52] These are questions about the nature of wrongs and the nature of punishment. While these new dataveillance technologies mean greater success in identifying, solving, and punishing wrongful conduct, such as terrorism, they are equally capable of identifying, solving, and punishing wrongful conduct of a more morally ambiguous nature. Consider, as an almost trivial example, the use of red light cameras in several major American cities. Before the development of this technology, drivers running red lights were identified only infrequently when they had the bad luck to run the light in the presence of a police officer. Now, with automated cameras, the rate of capturing wrongful red light runs is higher.[53] The same is increasingly true of a host of other offenses. Given the rate and scope of technological development, the trend will only continue. This change—the use of technology to make it more likely (if not certain) that violations of law will be observed—will work powerful effects on the deterrence component of law enforcement and, if properly applied, on criminal and espionage-type activity in cyberspace. We now calculate the optimal level of punishment by discounting the "real" punishment to account for the likelihood of getting caught. A ten-year sentence with a one-in-ten chance of capture arguably has an effective deterrent value of one year in prison. When the chance of capture increases, the effective deterrent does as well.

---

[50] *See* Helene Cooper, "Passport Files Of 3 Hopefuls Are Pried Into," *The New York Times,* Mar. 22, 2008, at A1.

[51] Two contract employees were fired by the State Department in the Obama case and a third was disciplined. *Id.* In the case of Joe Wurzelbacher ("Joe the Plumber"), whose tax records were disclosed, several Ohio state employees were identified and disciplined. *See* "Clerk Charged with Unlawful Search of Joe the Plumber," http://www.toledoonthemove.com/news/story.aspx?id=213580.

[52] I first discussed the ideas in this section with my friend and colleague Kim Taipale of the Center for Advanced Studies. *See also* K.A. Taipale, *Play Room in the National Security State* (unpublished manuscript, on file with the author) (Center for Advanced Studies Working Paper Series No. 05:0515) (technological changes are transforming criminal justice system from one based on punishment and deterrence to one based on ubiquitous preventative surveillance and control through system constraints).

[53] *See*, *e.g.*, Kevin Courtney, "Red Light Cameras Work, But Are Fines Too High?," *Napa Valley Register,*, Feb. 14, 2010, http://www.napavalleyregister.com/news/local/article_1fbc2456-1932-11df-b32f-001cc4c03286.html.

An interesting corollary to the development of new technologies is that they will, inevitably, require either a reduction in punishments across the board or a much better, and narrower, definition of "wrongful conduct." As technology trends towards near perfect enforcement, society will need to re-examine its definition of what constitutes a "wrong." To put it prosaically, in a world where we could identify every Senator who has illegally smoked a Cuban cigar or every individual who has exceeded the speed limit by the least amount, we might well need to change our definition of those acts as wrongful. Increasingly, we will need to consider how we can best enhance individual autonomy, and that may necessitate decreasing the sphere of governmental authority.

Thus, one of the unseen perils to dataveillance is not, as most privacy advocates suppose, the increased likelihood that the state will abuse its power by targeting for adverse consequence those who have committed no crime— for example, a person whose only act is to engage in political protest. The new structures and systems we are putting in place are likely to be capable of protecting against abuse. The real peril is that our conception of the state's ambit has grown so broad that the state may soon lawfully use its powers to target "wrongful" conduct that ought not, truly, to be deemed wrongful.

**Conclusion**

It will be a significant challenge to determine the right answers to many of the substantive questions I have posed. There will be substantial policy issues to resolve, for example, in determining what, if any, triggers might be created for denying an individual employment in a nuclear facility or refusing to let him board a plane. Yet these are the questions that must be answered. The improvements in computational power and data storage costs will not slow down, and we cannot expect to stop the deployment of new anonymity-invasive technology. Indeed, any effort to do so is doomed to failure before it has begun.

Therefore, rather than vainly trying to stop progress, or trying to fit the new technologies into old principles of privacy that no longer apply, we will need to go about the business of answering the hard policy questions. Instead of reflexively opposing technological change, a wiser strategy is to accept the change and work within it to channel change in beneficial ways.

This will require a rethinking of privacy—both a re-conception of what we think it means and a reconfiguration of how we think it is to be protected. It may be true that "privacy is dead," but for those who truly want to protect privacy, the motto should be: "Privacy is dead. Long live the new privacy."