INFORMATION SECURITY AND PRIVACY ADVISORY BOARD

Established by the Computer Security Act of 1987
[Amended by the Federal Information Security Management Act of 2002]

Toward A 21st Century Framework for Federal Government Privacy Policy

May 2009

INFORMATION SECURITY AND PRIVACY ADVISORY BOARD

Established by the Computer Security Act of 1987
[Amended by the Federal Information Security Management Act of 2002]

May 27, 2009

The Honorable Peter Orszag
Director
The Office of Management and Budget
725 17th Street, NW
Washington, DC 20503

Dear Mr. Orszag:

I am writing to you on behalf of the Information Security and Privacy Advisory Board (ISPAB). The ISPAB was originally created by the Computer Security Act of 1987 (P.L. 100-35) as the Computer System Security and Privacy Advisory Board, and amended by The E-Government Act of 2002, Title III, The Federal Information Security Management Act (FISMA) (P.L. 107-347). One of the statutory objectives of the Board is to identify emerging managerial, technical, administrative, and physical safeguard issues relative to information security and privacy.

Attached to this letter is a Board report that analyzes issues and makes recommendations around updating privacy law and policy in light of technological change. The Privacy Act of 1974 is the basis for much of the legal and policy framework by which the U.S. Government handles personal information. At the same time, vast changes in technology since 1974 have transformed how Federal agencies collect, use, and distribute information in major ways. While the fundamentals of the Act—the principles of fair information practices—remain relevant and current, the letter of the Act and related law and policy may not reflect the realities of current technologies and information systems and do not protect against many important threats to privacy. Moreover, new technologies, not covered by the Act, are generating new questions and concerns; and government use of private-sector databases now allows the collection and use of detailed personal information with little privacy protections. The attached report examines these issues, and is based on a record that has been developed through the Board' having heard from numerous panels of experts for several years. The Board provides analysis and makes recommendations for the Administration and Congress to consider.

We appreciate the opportunity to offer the Board's views on this critically important issue. Please let me know if the Board can answer any questions or take additional actions regarding privacy law and policy in light of technological change.

Sincerely,

Daniel J. Chenok
Information Security and Privacy Advisory Board Chairman

cc: Vivek Kundra
Administrator, Office of E-Government and Information Technology
and Federal Chief Information Officer

Kevin Neyland
Acting Administrator, Office of Information and Regulatory Affairs

Table of Contents

I. Summary

1. Executive Summary

Soon after passage of the Privacy Act of 1974, experts noted loopholes in the law. The Act's limitations have become more significant with the passage of time, as information technology has become more prevalent in the operation of government programs. And while the fundamentals of the Act—the principles of fair information practices—remain relevant and current, the letter of the Act and related law and policy do not reflect the realities of current technologies and do not protect against many important threats to privacy.

Inattention by policymakers to the underlying problems, and relatively little White House guidance, has meant that privacy policy is left to the individual agencies. There has been a lack of government-wide direction, and only a few privacy leaders in key agencies have been empowered by their internal leadership to fill the policy vacuum.

Moreover, new technologies not covered by the Act are generating new questions and concerns. For example, the Federal government has provided no guidance on technologies that allow civilian government agencies to track individuals and retain data about individuals by default. And government use of private-sector databases now allows the collection and use of detailed personal information with few privacy protections. Because little guidance has been provided to the agencies since the Privacy Act was enacted, agency policy and procedure have not adapted to technological change.

The Information Security and Privacy Advisory Board finds that the Privacy Act and related policy should be brought up to date. To begin to create a new framework to protect privacy, ISPAB makes the following recommendations:

- Amendments to the Privacy Act of 1974 and Section 208 of the E-Government Act of 2002 are needed to:
    - Improve Government privacy notices;
    - Update the definition of System of Records to cover relational and distributed systems based on government use, not holding, of records.
    - Clearly cover commercial data sources under both the Privacy Act and the E-Government Act.
- Government leadership on privacy must be improved.
    - OMB should hire a full-time Chief Privacy Officer with resources.
    - Privacy Act Guidance from OMB must be regularly updated.
    - Chief Privacy Officers should be hired at all "CFO agencies."
    - A Chief Privacy Officers' Council should be developed.
- Other changes in privacy policy are necessary

- OMB should update the federal government's cookie policy.
- OMB should issue privacy guidance on agency use of location information.
- OMB should work with US-CERT to create interagency information on data loss across the government
- There should be public reporting on use of Social Security Numbers

2. About ISPAB

The Information Security and Privacy Advisory Board (ISPAB) was originally created by the Computer Security Act of 1987 as the Computer System Security and Privacy Advisory Board. As a result of the E-Government Act of 2002, Title III, The Federal Information Security Management Act of 2002, the Board's name was changed and its mandate was amended.

The objectives of the ISPAB are as follows:

- Identify emerging managerial, technical, administrative, and physical safeguard issues relative to information security and privacy;
- Advise the National Institute of Standards and Technology (NIST), the Secretary of Commerce, and the Director of the Office of Management and Budget on information security and privacy issues pertaining to federal government information systems, including thorough review of proposed standards and guidelines developed by NIST.
- Annually report its findings to the Secretary of Commerce, the Director of the Office of Management and Budget, the Director of the National Security Agency and the appropriate committees of the Congress.

The Board's authority does not extend to private sector systems or to federal systems that process classified information.

The Board has examined issues around updating privacy law and policy for several years, having heard from numerous panels of experts. Information provided to the Board on the subject, as well as Board deliberations and general information about the Board, are available at the Board's website, http://csrc.nist.gov/groups/SMA/ispab.

This white paper is being released by the current Board, the membership if which is listed below. The paper was developed under the lead role played by Board member Ari Schwartz. In addition, the Board wishes to thank all of the experts who contributed to its content along the way. Special thanks are also due to members of the Board who served for much of the period during which this report was discussed, including Frank Reeder, Susan Landau and Phil Reitinger, and especially to Leslie Reis, who developed a foundational framework for the Board's thinking on this subject prior to the end of her tenure.

ISPAB Membership

Daniel Chenok (*Chairperson*)
Senior Vice President
Pragmatics

Jaren P. Doherty
Associate Deputy Assistant Secretary
Office of Cyber Security
Department of Veterans Affairs

Brian Gouker
Senior Advisor, Information Assurance Directorate
National Security Agency

Joseph A. Guirreri
Kforce

Rebecca C. Leng
Deputy Assistant Inspector General for Information Technology and Computer Security
U.S. Department of Transportation

F. Lynn McNulty
McNulty and Associates

Alexander L. Popowycz
Vice President
Fidelity Investments

Lisa Schlosser
U.S. Environmental Protection Agency (EPA)

Howard A. Schmidt
President and CEO
R & H Security Consulting LLC

Fred B. Schneider
Samuel B. Eckert Professor of Computer Science
Cornell University
Ari M. Schwartz
Vice President and COO
Center For Democracy and Technology

Peter J. Weinberger
Senior Software Engineer
Google, Inc.

II. Background

1. *Policy History*

In the early 1970s, the federal government was the worldwide leader in developing policies and best practices to protect the information held about citizens and other individuals. Those early efforts provided a firm basis for privacy in the federal government and remain a critical foundation for efforts to address new technologies. A review of U.S. federal government privacy history is necessary to help understand and identify where the gaps have arisen in current policy.[1] Important milestones include:

1. HEW Study on Fair Information Practices

In 1972, Elliot L. Richardson, then Secretary of the U.S. Department of Health, Education and Welfare (HEW), appointed an Advisory Committee on Automated Personal Data Systems to explore the impact of computerized record keeping on individuals. In the committee's report, published a year later, the Advisory Committee proposed a Code of Fair Information Practices (FIPS).[2] These principles formed the basis for all subsequent codes and laws related to information collection, especially the Privacy Act of 1974 and the OECD privacy guidelines.

2. Passage of the Privacy Act

The Privacy Act of 1974[3] provides the main controls within federal government on the collection, use, and disclosure of personally identifiable information. The law was designed to protect individuals from an increasingly powerful and potentially intrusive federal government.

The Privacy Act incorporates the Code of Fair Information Practices recommended by HEW, giving individuals certain rights with respect to the

[1] There are many privacy statutes and policies that place privacy requirements on the federal government in some way. We try to raise those that have a direct impact on technology and are necessary to understand the subsequent policy discussion and our recommendations.

[2] U.S. Dep't. of Health, Education and Welfare, Secretary's Advisory Committee on Automated Personal Data Systems, *Records, computers, and the Rights of Citizens*, viii. (1973). <http://aspe.dhhs.gov/datacncl/1973privacy/tocprefacemembers.htm>

[3] Privacy Act of 1974, Public Law No. 93-579, 88 Stat. 1897 (Dec. 31, 1974), codified in part at 5 U.S.C. § 552a.

federal government's collection, use, and dissemination of personal information. Privacy Act provisions require notice to and consent from individuals whenever the government collects and shares information about them, give citizens the right to see whatever information the government has about them, and hold government databases to certain accuracy standards. The Act ensured, through public Systems of Records Notices (SORNs) in the Federal Register, that the government would not be able to maintain secret databases on its citizens.

The Act also prohibits agencies from disclosing records to third parties or other agencies without the consent of the individual to whom the record pertains. There are several important exceptions where notice to the individual is not required.[4]

OMB was called upon to issue implementing guidance and issued a comprehensive document in July 1975, soon after the passage of the Act.[5] No

---

[4] 5 U.S.C. § 552a(b) — These include sharing information:
(1) to those officers and employees of the agency that maintains the record who have a need for the record in the performance of their duties; (2) as required to be disclosed under the Freedom of Information Act; (3) for a routine use, defined as "the use of such record for a purpose which is compatible with the purpose for which it was collected;" (4) to the Bureau of the Census for purposes of planning or carrying out a census or survey or related activity; (5) to a recipient who has provided the agency with advance adequate written assurance that the record will be used solely as a statistical research or reporting record, and the record is to be transferred in a form that is not individually identifiable; (6) to the National Archives and Records Administration as a record that has sufficient historical or other value to warrant its continued preservation by the United States Government, or for evaluation by the Archivist of the United States or the designee of the Archivist to determine whether the record has such value; (7) pursuant to a specific written request, to another agency or to an instrumentality of any governmental jurisdiction within or under the control of the United States for a civil or criminal law enforcement activity if the activity is authorized by law; (8) to a person pursuant to a showing of compelling circumstances affecting the health or safety of an individual; (9) to either House of Congress, or, to the extent of matter within its jurisdiction, any committee or subcommittee thereof, any joint committee of Congress or subcommittee of any such joint committee; (10) to the Comptroller General, or any of his authorized representatives, in the course of the performance of the duties of the General Accounting Office; (11) pursuant to the order of a court of competent jurisdiction; or (12) to a consumer reporting agency.

[5] Privacy Act Implementation, Guidelines and Responsibilities, 40 FR 28948, July 9, 1975 (July 8, 1975), http://www.whitehouse.gov/omb/inforeg/implementation_guidelines.pdf and

Administration has systematically updated the Privacy Act guidance since 1975.

It is important to note that the Privacy Act protections only apply to "systems of records," which are defined as databases where records are regularly located by a specific and unique identifier, such as a name or government issued ID number. Courts have since determined that even if a record is "locatable" by such an identifier and has been accessed irregularly by name, then it is not held in a system of records unless regularly and systematically retrieved by this same identifier.[6]

The Act also limited the use of the social security number (SSN) as an identifier. Older federal systems utilizing the SSN were not required to change, but new federal systems had to have explicit Congressional approval to use the SSN as an identifier.

3. Privacy Protection Study Commission Report

The Privacy Act created the Congressional Privacy Protection Study Commission (PPSC) to study privacy issues and recommend future legislation. The PPSC released its final report in July 1977, entitled "Personal Privacy in an Information Society."

In its main Privacy Act findings, the Commission strongly stated that the Act was too vague to meet its stated purposes. Specifically, the Commission found that:

- the definition of "systems of records" was too restrictive and that a broader definition incorporating how agencies accessed records was needed;
- the routine use exemption was unclear and was already being used in unintended and poorly disclosed ways; and
- the SORNs that were being published were unhelpful in informing the public about policies and practices.

However, the Commission's recommendations to address these problems were never passed by Congress nor addressed in any of OMB's future privacy guidance. Many current issues with the Privacy Act were identified by the PPSC in its 1977 report.[7]

Supplemental Guidance, http://www.whitehouse.gov/omb/inforeg/implementation1974.pdf

[6] Several cases, most notably Henke v. United States Department of Commerce, 83 F.3d 1453 (D.C. Cir. 1996) have emphasized the importance of this distinction.

[7] Personal Privacy in an Information Society, The Report of The Privacy Protection Study Commission, July, 1977; most of the report is accessible at < http://aspe.hhs.gov/datacncl/1977privacy/toc.htm> and < http://epic.org/privacy/ppsc1977report/>

4. OECD FIPs

In 1980, the Organization for Economic Co-operation and Development (OECD) adopted a series of guidelines designed to harmonize international privacy legislation without interrupting the free flow of information between borders.[8] The Fair Information Practices (FIPs), clearly based on the HEW principles, are broken down by OECD into eight principles that cover the collection, security, and primary and secondary uses of the data. These principles have become the baseline for evaluating privacy and data protection initiatives worldwide, and are:

- Collection Limitation Principle,
- Data Quality Principle,
- Purpose Specification Principle,
- Use Limitation Principle,
- Security Safeguards Principle,
- Openness Principle,
- Individual Participation Principle, and
- Accountability Principle.[9]

5. Computer Matching Act

The Computer Matching and Privacy Protection Act of 1988 (P.L. 100–503) added certain protections to records subject to the Privacy Act that are used in computer matching programs (ie, programs used to compare records in order to make a decision). This Act requires agencies sharing records with other agencies or non-Federal entities to enter into written agreements justifying and limiting matching, procedures for retention and destruction of data after matching, notification for individuals whose records are matched, and prohibitions on disclosure of records and the compilation of data. These agreements must be sent to specific congressional committees, available to the public upon request, and notice of programs must be published in the Federal Register.

Notably, the Computer Matching Act requires that individuals be allowed to refute information leading to actions taken as a result of this computer matching and establishes a Data Integrity Board in each agency.

---

[8] Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, OECD, 1980. <
http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html>

[9] http://webdomino1.oecd.org/horizontal/oecdacts.nsf/linkto/c%2880%2958

6. OMB Memo on cookies

OMB created a cookie policy for federal agency websites in 2000 in response to criticism of federal sites using cookies. OMB issued a memorandum[10] that states agency websites will not use cookies except in very limited circumstances. A September 5, 2000 letter from OMB to the Department of Commerce further clarified that "persistent" cookies, which remain on a user's computer for varying lengths of time (from hours to years), are not allowed unless these four specific conditions are met. "Session" cookies, which expire when the user exits the browser, are permitted.

In 2003, the E-Government Act privacy implementation guidance[11] expanded the types of technologies included under these policies to include any tracking technology that lasted longer than a user's one-time visit to the website.

Currently, in order to comply with the federal cookie guidelines, a government website must have a compelling need to gather the data through cookies, must provide users a clear and conspicuous notice of the use of cookies, must have a clear privacy policy explaining the collection of information through cookies and privacy safeguards for handling that information, and must have personal approval from the agency head, or a delegate who reports directly to them.

7. Privacy Impact Assessments

The E-Government Act of 2002 included important updates to government privacy policy, supplementing protections of the Privacy Act. Section 208 of the E-Government Act requires that agencies post on their Web sites privacy notices about their information collection practices. The Act also requires agencies to conduct privacy impact assessments (PIAs). Specifically, PIAs must be completed before developing or procuring new technology that collects, maintains, or disseminates personal information and before initiating new collections of personally identifiable information. Under the law, the PIAs are public documents and are supposed to contain a description of the project, a risk assessment, a discussion of potential threats to privacy, and ways to mitigate those risks. OMB issued guidance in 2003, as required by the Act, to help implement the Act and also used the opportunity to clarify government-wide privacy policy on several non-Privacy Act related issues, such as the use of tracking technologies on the Internet.

---

[10] OMB Memorandum 00-13, "Privacy Policies and Data Collection on Federal Web Sites" http://www.whitehouse.gov/omb/memoranda/m00-13.html

[11] OMB Memorandum 03-18, "Implementation Guidance for the E-Government Act of 2002" http://www.whitehouse.gov/omb/memoranda/m03-18.html

8. OMB Memo on data breach reporting

In 2007, OMB issued guidance on data breach and protecting information held by agencies. Specifically, OMB instructed agencies to reduce the risk of data breaches by reducing the amount of information retained by the agency, limiting access to that information, and protecting the information using access controls (such as authentication and encryption) to make records unusable in the event of a data breach. This guidance also instructed agencies to develop policies for notification of a data breach, whether the information is about federal employees or others.

9. DHS FIPs Framework

In December 2008, the Department of Homeland Security issued guidance on its interpretation of the Fair Information Practices (FIPs).[12] The DHS framework offers significant updates to the OECD FIPs and has already received attention from privacy professionals and others as a new model. The Framework makes progress in clearly stating many modernized versions of FIPs and tying them directly to the Privacy Act for the agencies use. For example, included in this framework is the concept of data minimization — "directly relevant and necessary to accomplish the specified purpose(s) and only retain PII for as long as is necessary to fulfill the specified purpose(s)." In the past, the more limited concepts of "use limitation" and "collection limitation" were pointed to explain the Privacy Act provision that makes clear that an agency should "maintain in its records only such information about an individual as is relevant and necessary to accomplish a purpose of the agency required to be accomplished by statute or by Executive order of the President."[13] This new definition helps clearly express the Department's, and the federal government's, basic principles for privacy to the world.

2. GAO Findings on the Privacy Act

The Government Accountability Office (GAO) has issued a number of reports on the state of the Privacy Act in light of technological and policy changes.[14] In

[12] DHS Privacy Policy Guidance Memorandum 2008-01, The Fair Information Practice Principles: Framework for Privacy Policy at the Department of Homeland Security, December 29, 2008, <
http://www.dhs.gov/xlibrary/assets/privacy/privacy_policyguide_2008-01.pdf>

[13] 5 U.S.C. § 552a(e)(1)

[14] GAO issued multiple reports over this time. We cite those that directly influenced this report. The most recent of these — GAO, Privacy: Congress Should Consider Alternatives for Strengthening Protection of Personally Identifiable Information , GAO-08-795T (Washington, D.C.: June 18, 2008) —summarizes much of GAO's past work in this area.

2003, at the request of Congress, GAO analyzed the laws and guidance of the Privacy Act and consulted with federal agencies to determine how agencies have implemented the Act. The report, entitled "Privacy Act: OMB Leadership Needed to Improve Agency Compliance," highlighted the fact that there has been almost no detailed Privacy Act guidance from OMB since 1975, and that since 2000 OMB has offered little or no support to agencies with questions about implementation of the Act.[15] A 2004 report on data mining practices suggested widespread practices, while clearly covered in the spirit of the law, that it is not clear whether these practices were covered by its actual terms.[16]

Finally, just last year, GAO released a detailed review of the Privacy Act, finding that problems go beyond just implementation. The law itself, GAO concluded, is simply not sufficient to protect the privacy of individuals today.[17] This report made clear that the long-standing problems in the Privacy Act have not been mitigated.

The 2008 GAO report made three primary findings, which were accompanied by detailed proposals for addressing the problems:

1) "Applying privacy protections consistently to all federal collection and use of personal information."

The GAO found that the definition of a "system of records" is not universally applicable to the types of personally identifiable information collected by the government. The GAO recommended revising this definition to cover all personally identifiably information that is collected by the federal government.

2) "Ensuring that collection and use of personally identifiable information is limited to a stated purpose."

The GAO found that the current privacy regime does not adequately control collection and use of personally identifiable information. In response, the GAO recommended that the law be amended to require agencies to justify collection of information and to justify the use or sharing of personally identifiably information.

3) "Establishing effective mechanisms for informing the public about privacy protections."

---

[15] GAO, Privacy Act: OMB Leadership Needed to Improve Agency Compliance, GAO-03-304
(Washington, D.C.: June 30, 2003).

[16] GAO, Data Mining: Federal Efforts Cover a Wide Range of Uses, GAO-04-548 (Washington, D.C.: May 4, 2004).

[17] GAO, Privacy: Congress Should Consider Alternatives for Strengthening Protection of Personally Identifiable Information , GAO-08-795T (Washington, D.C.: June 18, 2008).

The GAO found that current methods used to inform the public about policies and practices around government collections of information are ineffective. Specifically, Privacy Act notices are hard to understand and difficult to find. The GAO recommended the use of layered notices, in which the most important facts are presented to the user to begin with, followed by denser and more esoteric information as the user digs deeper. The GAO also recommended publishing these sorts of notices at a central, easy to access location on the Web.

III. Changes in Technology and their Impact on Federal Privacy Policy

New information technologies have changed how data can be collected, used, and shared. This section focuses on some key technologies that have had such an impact.

1. Relevant changes in database technology and information access since 1974

Even as it was being adopted, the Privacy Act was being mooted by technological changes. In the early 1970s, databases were generally held in centralized locations[18] on "flat file" databases[19] that could only be searched through a specific field; hence, that is how the Privacy Act defined a covered "system of records." However, from the mid 1970s into the 1980s, two major developments occurred: a move from large centralized machines to personal computers and a move to relational databases. Relational databases, which group data using common attributes, revolutionized how information can be processed and retrieved.[20] Database managers no longer need a central numbering system or a unique

[18] Databases kept in these central locations were accessed through "dumb terminals" in locations other than the location housing the data. Users would query the database through the terminal in front of them, but all processing and lookup happened in the centralized location where the data was stored. This centralization was necessary due to the relatively small processing power of computers and the small number of database functions available.

[19] Flat file databases are encoded as a plain text file and thus have very little structure that could enable what we now think of as database retrieval functions. Typically, flat file databases contain one record per line, and commas or other 'special' characters delimit fields. Thus, it is possible to write out a flat file database on a sheet of paper. Flat file databases require interpretation through a utility that enables database functions. For more information, see http://www.pcmag.com/encyclopedia_term/0%2C2542%2Ct%3Dflat+file&i%3D43286%2C00.asp and http://en.wikipedia.org/wiki/Flat_file_database

[20] http://www.pcmag.com/encyclopedia_term/0%2C2542%2Ct%3Drelational+database&i%3D50369%2C00.asp

identifier to compare sets of records or sort data. More recently, the Internet has promoted the growth of fully distributed databases.[21] These developments have resulted in a much different computing environment. Today, it is rare that personal information is stored on a single central computer and is only searched by referring to a single discrete identifier. Policy written for the era of flat files has confused and frustrated those who would like to follow the law, especially since there has been no government-wide guidance on how to apply the Act's older terms and assumptions to today's environment.

2. *Growth of technologies that track individuals*

In recent years, there has been dramatic growth in the number and types of technologies that can track individuals—both technologies explicitly designed for tracking and those created for other purposes that can be used to track individuals. These newer technologies also include those that do not necessarily use personal information but that may have an impact on personal privacy depending on how they are used. And they are generally not explicitly covered by federal policies. Some specific examples include:

1. Aggregation of non-identified data, including commercial search, cookies and log information, to create identifiable records

Most traditional privacy concerns have focused on *personally identifying information (PII)*—name, address, phone number, Social Security number and other government-issued identification numbers, email address, and financial identifiers such as credit card numbers—that may be tied back to a specific person. The coverage of the Privacy Act, as noted above, is limited to databases organized according to such identifiers. Increasingly, it has become clear that significant privacy concerns can arise from non-identifying personal information.[22]

a. Search and Use of IP Addresses

For example, the enormous growth of the Web has brought search engines and other tools that help users find information. Indeed search engines have become essential for all Internet users. For a variety of reasons, search engines collect information on individual searches and the use of search results (e.g., for advertising and to improve search engine effectiveness). These stored results are usually not tied to a user's name nor email address, but to the Internet Protocol (IP) address of the user's

---

21 http://www.pcmag.com/encyclopedia_term/0,2542,t=distributed+database&i=41561,00.asp

22 The ISPAB addressed this issue in a letter to OMB last year NIST Information Security and Privacy Advisory Board December 10 letter to Jim Nussle http://csrc.nist.gov/groups/SMA/ispab/documents/correspondence/ISPAB_Einstein-letter.pdf.

machine and a cookie ID number or a related ID number created and placed on a user's computer by the search engine.

This information is not as anonymous as some would assume. In August 2006, AOL publicly released "anonymized" log files containing twenty million search queries for over 650,000 users over a 3-month period; the data included a unique identifier for each user but did not include anything that would traditionally have been considered PII. Nevertheless, several researchers were easily able to identify individuals based on these "anonymized" records.[23] While it had been proved years earlier that non-identifying personal information (such as birth date, zip code, and gender) that can be gleaned or inferred from query logs can be used to link queries to an individual if combined with other publicly available data, such as census or voter registration databases,[24] the data available solely in regular search queries or online browsing activity was not widely understood to be a major privacy concern until the AOL breach case. Today it is clear that logs of online behavior may provide insight into many activities, hence can invade privacy.

More recently, the Federal Trade Commission (FTC) raised similar concerns in the context of online behavioral advertising. Finding that "the traditional notion of what constitutes PII versus non-PII is becoming less and less meaningful and should not, by itself, determine the protections provided for consumer data,"[25] the FTC staff cited 5 reasons for heightened concern:

1) "[D]epending on the way information is collected and stored, it may be possible to link or merge non-PII with PII;"
2) It is becoming "easier to identify an individual consumer based on information traditionally considered to be non-PII;"
3) "[E]ven where certain items of information are anonymous by themselves, they can become identifiable when combined and linked by a common identifier;"
4) Combining detailed sets of non-identifiable data, such as research on a medical condition and prescription drugs, could "constitute a

[23] Micheal Barbaro, Tom Zeller Jr. and Saul Hansell, "A Face Is Exposed for AOL Searcher No. 4417749" New York Times, August 9, 2006. The Times was able to identify several individuals and the Electronic Frontier Foundation was able to identify many others.

[24] LaTanya Sweeney, "Uniqueness of Simple Demographics in the U.S. Population," LIDAP-WP4. Carnegie Mellon University, Laboratory for International Data Privacy, 2000.

[25] "FTC Staff Report: Self-Regulatory Principles For Online Behavioral Advertising" February 2009 http://www.ftc.gov/os/2009/02/P085400behavadreport.pdf p.21.

highly detailed and sensitive profile that is potentially traceable to the consumer;" and

5) Research shows that "consumers are concerned about the collection of their data online, regardless of whether the information is characterized as PII or non-PII."

Based on this review, the Commission determined that policies were needed to protect non-PII in the behavioral advertising context.[26]

While the FTC's detailed description of concern was limited to the behavioral advertising context, the analysis holds true in other contexts as well. In fact, in 2007, the European Union's Article 29 Working Party detailed an opinion on the concept of "personal information," suggesting that information that could eventually be used to identify individuals, such as an IP address, should be treated as identifiable in most instances.[27]

Based on public concerns about how search log information could be used and a separate regulatory threat from the Article 29 Working Party,[28] search companies began to limit the retention of information in their logs.[29] These efforts have helped to minimize the amount of data stored and used, and they have also put a spotlight on the privacy implications of information that has not traditionally been considered personally identifiable.

Undeniably, it is difficult to achieve a good policy balance when addressing the use of non-personal information. If policymakers go too far, then all information of any kind could be deemed to have privacy sensitivity. If policymakers do not go far enough and only cover PII, then this category of potentially revealing data will be inadequately protected. Some experts (see, for example, Nissenbaum[30]) have begun suggesting an approach that focuses more on specific uses of any kind of data than on the data itself, but this field of study is still young.

---

[26] FTC Staff Report: Self-Regulatory Principles For Online Behavioral Advertising" February 2009 http://www.ftc.gov/os/2009/02/P085400behavadreport.pdf pp. 21-25.

[27] Article 29 Data Protection Working Party, "Opinion 4/2007 on the concept of personal data" June 2007 http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf

[28] Article 29 Data Protection Working Party, "Opinion 1/2008 on data protection issues related to search engines" January 2008 http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2008/wp148_en.pdf

[29] Search Privacy Practices: A Work In Progress, Center for Democracy and Technology, August 2007, http://www.cdt.org/privacy/20070808searchprivacy.pdf

[30] Helen Nissenbaum. Privacy as Contextual Integrity. *Washington Law Review*, 79(1): 119-570, 2002

b. Use of Cookies

The federal government has had its own sets of concerns with non-PII. In fact, concerns over the federal government's use of log information and the logging practices of agency partners go back even further than AOL's posting of search results. In 2000, the Office of National Drug Control Policy sparked controversy by partnering with DoubleClick to place tracking cookies on the ONDCP Web site as part of an advertising campaign.[31] This incident led the White House to explicitly ban the use of cookies in most circumstances.[32] The policy was loosened slightly in 2003 to target only persistent cookies but also expanded to cover Web beacons and other third-party tracking identifiers.[33] Federal Webmasters have complained that the current policy inhibits their ability to improve services.[34] Recently, the Obama White House's move to exempt certain kinds of cookies raised this issue anew.[35]

Unquestionably, finding a balancing point in policy between privacy concerns and the ability to analyze and improve service delivery remains a difficult goal in both the private and government sectors. However, now that greater attention has been paid to the potential privacy risk from use of IP addresses, cookies and similar data, perhaps finding a proper balance can become a reality.

2. Social Networking

The biggest change in Internet technology in the past three years has been the growth in Web 2.0 interactive tools. Social networks such as Facebook, MySpace and Twitter have changed the way individuals think about how they communicate online. Young people, in particular, expect to interact with the Web sites they use and to create content themselves rather than merely use content created by others.[36] At the same time, however, this new relationship with the Internet is resulting in the transfer of personal information to "the cloud" at an unprecedented rate.

---

[31] http://shns.scripps.com/shns/story.cfm?pk=COOKIES-06-20-00&cat=AN

[32] Privacy Policies and Data Collection on Federal Web Sites, http://www.whitehouse.gov/omb/memoranda/m00-13.html

[33] OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002, http://www.whitehouse.gov/omb/memoranda_m03-22/

[34] Bev Godwin, et al "Social Media and the Federal Government: Perceived and Real Barriers and Potential Solutions" December 24, 2009 http://techpresident.com/blog-entry/social-media-and-federal-government-perceived-and-real-barriers-and-potential-solutions

[35] Chris Soghoian,"White House acts to limit YouTube cookie tracking" January 23, 2009 http://news.cnet.com/8301-13739_3-10148844-46.html

[36] http://www.businessweek.com/magazine/content/07_24/b4038405.htm

Social networks bring together groups of users based on shared interests, geography, or other attributes. These networks allow users to make connections, share information, or communicate by storing information on the computers of the service provider. Social networking sites have grown exponentially, attracting a young and engaged audience who often share information freely with each other. Most social networks give users privacy controls—though not all users take advantage of them. Sometimes, this information is accessible to others who are not intended recipients.

Millions of Americans use social networks and share large amounts of personal information. Increasingly, social networking services are being adopted by older users for both personal and professional reasons. In the United States, MySpace and Facebook are the largest social networks. According to a 2009 Pew Internet survey, 35% of American adults have a profile on at least one social network, and 65% of American teens have social networking profiles.[37]

The Obama administration has actively been pursuing activities to use social networking and further increase transparency and participation in government using technology. Staff from OSTP, GSA and OMB are working on an Open Government Directive to instruct agencies and executive departments how to implement the principles of transparency, participation, and collaboration- and these will inevitably involve the use of technology in ways that are not anticipated in the Privacy Act.

Even before the completion of the Open Government Directive, federal government agencies have begun to engage with users through social networks and other Web 2.0 services.[38] NIH has made forays into Second Life, for example, while other agencies have established organizational pages on Facebook. GSA has recently negotiated alternate contracts for government use of popular social networking sites, and is continuing to negotiate access to services for government. The government can engage in these services while protecting security and privacy, but the complexity of the flow of information in such environments will require changes in current policy.

Social networking poses a range of privacy issues. These concerns range from use by law enforcement to use by third party data brokers who may begin to incorporate information from social networking sites into their

---

37 http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Adult_social_networking_data_memo_FINAL.pdf.pdf

38 Doug Beizer, "Twitter, blogs and other Web 2.0 tools revolutionize government business" Federal Computer Week, March 6, 2009.

profiles. Information is voluntarily shared by users, but they may not anticipate government access of or use of the information to make decisions about them.

3. Location data/Geocoding

Mobile device technology has undergone rapid revolution in recent years, with more and more sophisticated features available on smaller and smaller devices. Location determination technologies—including GPS chips, cellular tower and WiFi base station location capability, and other technologies—have been at the forefront of this trend. An ever-increasing variety of mobile handsets and laptop computers can be geo-located more precisely than ever before. The ubiquity of increasingly powerful mobile devices has already spawned the first generation of location-based services and applications that make use of an individual's geographic position.

While location technologies offer great utility to the individuals and organizations that use them, the trail of information left behind raises concerns about privacy and security. Because individuals normally carry their mobile devices with them, location data can be compiled to form a comprehensive record of an individual's movements and activities. While other kinds of data—an individual's medical record or tax return, for instance—could be considered more sensitive than location information in certain contexts, these more traditional kinds of data provide only snapshots of an individual's activities at discrete moments in time or within discrete aspects of their lives. Location data, on the other hand, may be collected everywhere and at any time, often without user interaction, and it may potentially describe both what a person is doing and where he or she is doing it. Location-based services may allow for amassing such data points about an individual's every movement, potentially supporting the creation of richly detailed profiles of individual behavior.

Mobile phones, probably the primary source of location data and location-based tools today, are extremely personal devices precisely because users carry them everywhere they go. Unlike a desktop computer, which may be shared among multiple family members or visitors to a library or Internet cafe, most mobile handsets are carried everywhere and used only by a single individual. Even more so than laptops, mobile handsets have the potential to become personal homing beacons. For example, the location information from a cell phone may reveal the fact that an individual was in a particular

medical clinic or government building, for example, implying information about the individual that was not meant to be shared.

Often the two most tangible privacy harms cited for any piece of revealing information are embarrassment or identity theft, but direct access to location information through mobile devices adds an immediate safety concerns for specific at-risk individuals. The spread of location-based services may also increase the risks of stalking and domestic violence if perpetrators are able to use (or abuse) location-based services to gain access to location information about their victims. Location information also raises enormous child safety concerns as more and more children carry mobile devices.

As the accuracy of location data improves and the expense of calculating and obtaining it declines, location information may well come to pervade the Internet experience, serving as a core component of local search, location-based social networking, emergency services, and as-yet unimagined applications. Federal agencies may find all manner of novel uses for location information, from helping citizens locate the nearest national park to geo-targeting public safety alerts to specific locales.

Given the sensitivity of location information and the potential that location-based services create for the physical tracking of individuals, federal agencies seeking to collect location data must address privacy in the very design of services and applications or risk losing the trust of the very citizens they aim to engage.

The creation of databases of location information within the federal government would pose the possibility of "mission creep." Current government privacy laws do not provide sufficient limits on how location information may be used once it is collected by a federal agency. While this problem is not limited to location information, the possibility that some agencies may soon be collecting location data highlights the privacy risks posed by the fact that the Privacy Act and other federal guidance on the issue do not provide strong prohibitions against such secondary uses of information.

3. Government Use of Commercial Databases

The government not only collects personally identifiable information directly, it also buys information from commercial entities. Commercial vendors draw much of this information from public records at courthouses and other government agencies. The companies, sometimes known as data brokers, provide a valuable service to the private and government sectors

alike by aggregating and categorizing this information. Commercial data services companies also compile personally identifiable information that is not publicly available. This non-public, but commercially available data includes, for example, credit reporting information. Depending on context, it may also include a broad range of other data generated by individuals in the course of commercial transactions, online and off. Government agencies are heavily dependent on commercial data services. The GAO estimates that these contracts are valued at over $30 million at DHS alone.

While data brokers provide important services to the government and the private sector, the collection and aggregation of personally identifiable information also raises privacy issues and concerns about the accuracy, reliability and security of this information. Security breaches at all of the major data brokers have prompted calls for examination of security standards for this evolving industry.

However, the Privacy Act does not adequately cover government use of commercially-compiled databases of personal information. The rules about the federal government's use of commercial databases, and even use of information gleaned from commercial search engines, have been vague and sometimes non-existent. The Privacy Act's protections only apply to federally controlled systems of records,[39] which means the government may be able to bypass the protections of the Privacy Act by accessing existing private sector databases and searches, rather than collecting the information itself.

Subsection (m) of the Privacy Act covers government contractors. It was designed to ensure that an agency could not simply contract away its responsibilities for privacy protection under the Act. Subsection (m) simply states that, when an agency provides by contract for the operation on behalf of the agency of a system of records to accomplish an agency function, the agency shall cause the Privacy Act to be applied to such system. Similarly, all employees of such a contractor are bound by the Act to the same extent that federal employees would be. However, Subsection (m) does not adequately protect personal information in commercial databases that are merely being used by the government.

Situations involving Subsection (m) generally can be analyzed under four categories:

1. Private Collection Under Government Contract—The Privacy Act as currently written clearly applies when the government contracts with

[39] The term "system of records" is defined as "a group of any records under the control of any agency from which information is retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual." 5 U.S.C. § 552a(a).

a commercial entity to collect, maintain or analyze PII for use in performing a government function or program. The fact that the data is held by the commercial entity, and even the fact that no data ever enters government computers, makes no difference: all Privacy Act principles apply to the data in the private entity's computers that was collected at the behest of the government.

2. Merging of Private Sector Data—The Privacy Act does not clearly apply when commercial data is brought into government databases. According to most readings of subsection (m) of the Act, a new System of Records Notice (SORN) should be issued whenever contractor databases containing private sector data are used to augment existing systems of records housed by the government or its contractors.[40]
3. Receipt of Commercial Data—The Privacy Act as written does not clearly apply when PII is transferred to the government or its contractors from the private sector. However, there seems to be a lack of clarity about this issue. Under the Act, as narrowly interpreted, no covered "system of records" exists unless the identifiable information is not just "searchable" by name or other identifier but is actually searched by such means on multiple occasions. For example, the DHS Inspector General examined cases where commercial data on millions of individuals was appended to passenger flight records from airlines and held by a government contractor or by the government itself. The IG said that the Privacy Act was not violated because "the airline passenger records were not maintained in such a way as to have required TSA to publish a Privacy Act system of records notice," [41] presumably because data was not regularly searched on the basis of name. GAO disagreed and suggested that the Privacy Act may have been violated and the DHS Chief Privacy Officer ultimately agreed that the agency did, in fact, violate the Privacy Act.[42] The fact that the question of coverage arose at all highlights the need to clarify the Act.

---

[40] GAO, "Privacy: Government Use of Data from Information Resellers Could Include Better Protections," GAO-08-543T March 11, 2008 http://www.gao.gov/products/GAO-08-543T

[41] "Review of the Transportation Security Administration's Role in the Use and Dissemination of Airline Passenger Data," (Redacted), OIG-05-12, March 2005 http://www.dhs.gov/dhspublic/interweb/assetlibrary/OIGr-05-12_Mar05.pdf, at p. 45.

[42] GAO, "Aviation Security: Transportation Security Administration Did Not Fully Disclose Uses of Personal Information during Secure Flight Program Testing in Initial Privacy Notices, but Has Recently Taken Steps to More Fully Inform the Public" Memo to Congressional Committees, July 22, 2005, http://www.gao.gov/new.items/d05864r.pdf,
CDT Policy Post, "JetBlue Case," Volume 9, Number 20, October 17, 2003, http://www.cdt.org/publications/pp_9.20.shtml.
Privacy Office, Department of Homeland Security, "Secure Flight Report," December,

4. Direct Use of Private Sector Data Without Merging—The greatest area of uncertainty concerns whether the Privacy Act applies to commercial databases used by the government when the database a) was not created at the government's behest; b) remains in the control of the contractor or other commercial entity; and c) is queried by the government remotely. In this case, a government agency simply pings information from a commercial database and does not incorporate the data, but may use the information to make decisions about an individual.

   Two major sources for this type of use today are:
   - The regular use of non-contractor data sources (ie commercial search): the proliferation of commercial services, both free and for pay, has made information more easily accessible to everyone. Many human resource departments make it a policy to search public websites for information about potential hires, and many more will search for that person with or without a policy in place. The information from these searches is collected without notice or consent from the person to whom the information pertains, and without methods for notice or redress.
   - Occasional use of search for data gathering – when a source of information (regardless of its owner) is searched only occasionally, the Privacy Act does not protect the information used if it is not incorporated into an existing system of records. The Privacy Act did not anticipate the easy access of distant databases, and therefore these occasional searches bear information outside the definitions of the Privacy Act.

   Agencies seem confused by these different situations and there is concern that agency officials and government contractors are using this confusion to ignore or subvert the Privacy Act. Specifically, most agencies interpret the use of third party data that is outside a system to be outside the scope of the Privacy Act because no government system of records is involved. As there is no policy in place to as to how that information can be used or to address what protections this information should be given, agencies and staff do not have adequate guidance.

2006/ http://www.cdt.org/security/20061222secure.pdf,

In addition to concerns about the reach of the Privacy Act, there are concerns about Section 208 of the E-Government Act, which requires agencies to conduct Privacy Impact Assessments; as currently interpreted, Section 208 does not adequately address irregular government use of data compiled by the private sector.[43] OMB guidelines allow agencies to exempt the government's use of private sector databases from the requirement to conduct PIAs when the commercial data is not "systematically incorporated" into existing databases. This raises concerns about how to deal with information held by third parties. Companies that provide private sector data to the government have a range of security and privacy practices. Government agencies should use the PIA process to take those issues into account when making decisions about the use of commercial data.

Some agencies are already requiring PIAs for uses of commercial data even when the data is not integrated into existing databases despite OMB's guidance. In 2006, the GAO recommended that OMB revise its guidance to clarify the applicability of requirements for PIAs with respect to agency use of data obtained from commercial re-sellers. However, OMB did not address that recommendation and openly disagreed with it in House Oversight and Government Affairs Committee testimony in 2008.[44]

4. Data Mining

Data mining, broadly defined, encompasses activities designed to analyze large data sets and extract useful information, especially by uncovering patterns and relationships in the data that may not be apparent on first glance. Data mining is becoming easier and more common due to new technologies; most database software, even modern spreadsheets, come with data analysis capabilities. Data mining has been used or proposed for a variety of purposes, from detecting fraud to financial analysis to catching terrorists. Some instances of data mining fall within the definition of "computer matching" in the Privacy Act, which is the computerized comparison of two or more systems of records (or a system of records compared with non-Federal records) to determine eligibility for or compliance with the requirements of a federal benefits program. However, the Act is increasingly inadequate to address privacy in the modern data mining and data sharing environment.

---

[43] GAO, "Privacy: Government Use of Data from Information Resellers Could Include Better Protections," GAO-08-543T March 11, 2008 http://www.gao.gov/products/GAO-08-543T

[44] Statement of Karen Evans, Administrator of Electronic Government and Information Technology, OMB, before the House Subcommittee on Information Policy, Census and the National Archives of the Committee on Oversight and Government Reform, March 11, 2008 http://informationpolicy.oversight.house.gov/documents/20080318172705.pdf

In a 2004 report,[45] the GAO examined 199 data mining initiatives, of which 122 used personal information. This included 36 data mining efforts that included information from private, third-party sources. This same report noted that data mining was being increasingly used in efforts to detect threats or decide who may be a terrorist. A more recent CRS report says that data mining is one of the main techniques used in homeland security initiatives.[46]

Government use of data mining—whether on the government's own databases or using data from third parties—can impact privacy in several ways. In particular, poorly designed data mining allows inferences to be made about individuals based on relatively insignificant information. For example, an investigator could determine that if two terrorist suspects call the same number three times in the same month, it may —on its face — seem reasonable to want to determine others that have called this same number multiple times. However, if this number were a popular pizza delivery company in a major city, the government could very well be wasting its investigatory resources while invading the privacy of innocent Americans.

A recent National Research Council study on data mining, behavioral surveillance, and privacy, emphasized the importance of using effectiveness to judge whether a program should go forward. The study also recommended that a program be reevaluated for effectiveness every time that program changes in order to avoid "mission creep."[47]

Data mining techniques represent a fundamental change in the way the government accesses and uses data. In the past, the government collected and processed data on one person at a time (i.e., with particularity), either in the course of administering a government program or where there was some suspicion that a person was engaged in fraud, criminal conduct, terrorism or intelligence activity. The government was authorized to keep this data for long periods of time, and to retrieve, share and analyze it for compatible purposes without serious controls. New techniques like data mining undermine these protections as the government analyzes information en masse.

---

[45] GAO, Data Mining: Federal Efforts Cover a Wide Range of Uses, GAO-04-548 (Washington, D.C.: May 4, 2004). http://www.gao.gov/new.items/d04548.pdf

[46] Data Mining and Homeland Security: an Overview, August 27, 2008, http://assets.opencrs.com/rpts/RL31798_20080827.pdf

[47] Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals, National Research Council, *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*, National Academies Press, 2008.

The power of data mining to analyze information and produce inferences about individuals demands equally powerful privacy protections. Currently, however, the Privacy Act exempts information used for law enforcement from restrictions on computer matching programs, and the definition of computer matching may be so narrow as to exclude important data mining techniques.

5. *Increase in capability and portability of smaller and less expensive storage devices*

One major change readily evident in the computing over the past 35 years is the massive increase in the capacity and speed of computing power, while at the same time there have been equally dramatic innovations in storage technology enabling the retention of much larger amounts of data at lower costs on smaller and smaller devices. A NIST study has pointed out the nation's digital storage industry[48]—makers of the drives, tapes and other gear that have become the archives and the retrieval tools of the information age—has been doubling storage capacity about every 18 months.

The size of these computers continues to get smaller and smaller. Since the introduction of disk-based hard drives in 1956, the density of information it can record has increased 50-million times over, from 5MB to over 100 gigabytes[49]. As storage technologies developed, they also became exponentially cheaper. In 1960, this 5MB disk storage cost $50,000, but today storage of five megabytes of information costs around a nickel.[50] Similarly, storage of massive amounts of information can now be accomplished on a small stick of memory the size of a pen cap. This can easily be carried from place to place, and can easily be misplaced.

These changes in computing clearly offer the advantage of increased productivity. One unintended consequence, however, is that the personal information stored on these laptops and thumb drives is now much more portable and these devices and the increased information can be more easily lost or stolen. In 1974, when the Privacy Act was passed, no one was walking out the door of a government agency with personal data on 50 million people unnoticed. Today, this information can be copied and placed in someone's pocket. In fact, many data breaches and data loss occur because of lost and stolen laptops and storage devices.

---

[48] David Austin and Molly Macauley, "Estimating Future Consumer Benefits from ATP-Funded Innovation: The Case of Digital Data Storage", http://www.atp.nist.gov/eao/gcr_790.pdf, April 2000

[49] Chip Walters, "Kryder's Law," Scientific American, August 2005, http://www.sciam.com/article.cfm?id=kryders-law&ref=sciam

[50] http://www.alts.net/ns1625/winchest.html

6. *Distributed Computing*

Distributed computing, often called cloud computing, incorporates technology advances to store or process information at a location other than the local computer. Typically, these services are third parties and not operated by the government. When the Privacy Act was written, it was assumed that data used by federal agencies would also be held and processed on equipment within those agencies.

Distributed computing does not allow the user to personally secure the data and ensure its privacy rather it relies on the security of the vendor. The Privacy Act definitions make it clear that a system of records includes any information that is maintained by the agency or its contractor, which should cover any Privacy Act system in the cloud. However, we have seen that agencies have been confused about ambiguity over who is responsible for and legally holds records in other contracting relationships and cloud computing could create greater confusion. Discussion of the governments developed a central 'cloud computing' architecture for the use by agencies may alleviate concern over corporate access to these records, but may make internal sharing of these records easier in ways that may not be covered by existing policy.

7. *Authentication*

Authentication systems can play an important role in online service delivery and security; interest in authentication has increased dramatically as fraud and security concerns have grown. New technologies for authentication could make possible greater realization of the Internet's potential by making online transactions more seamless, tying together information on multiple devices, enabling yet unimagined services and taking us a few steps closer to a pervasive computing environment. Specifically in the context of e-government, there has been interest in the development of authentication systems to enhance delivery of government services online.

The CSIS Commission on Cybersecurity for the 44th President specifically focused on identity management as a major recommendation to help defend cyberspace from attack.[51] The Commission called for government-issued credentials to be used for critical online activities. However, ongoing discussions about government use of authentication systems raise concerns about government use of personal information and the creation of a centralized identity system. To mitigate these risks, as the Commission calls for, it is essential that authentication systems be designed to support effective privacy practices and offer individuals greater control over their personal information. Widespread adoption of authentication technologies will occur only if individuals trust that strong privacy

[51] http://www.csis.org/media/csis/pubs/081208_securingcyberspace_44.pdf

and security protections have been built into authentication systems themselves. As experts have noted, developing these protections is not an easy task.[52]

Fortunately, some work has begun to identify issues. In 2003, the Authentication Privacy Principles Working Group (APPWG) at the Center for Democracy and Technology issued a series of high-level privacy principles for authentication systems, which included calls to provide a diversity of authentication services and only use individual authentication where appropriate.[53] In 2004, OMB issued a memo providing "E-Authentication Guidance for Agencies;" the memo supported the APPWG principles and identified "Risk Levels and Risk Assessments" for agency consideration in adopting authentication technologies.[54]

In some ways, however, the hard work of balancing privacy and authentication as it relates to government is still to be completed. As the National Academies' Committee on Authentication Technologies and their Privacy Implications pointed out, government plays multiple roles as regulator, issuer, and relying party in the authentication process.[55] As government agencies make greater use of authentication technology, there will be a greater need for attention to privacy policy for such technology.

IV. Issues in Federal Privacy Policy Raised by Technological and Related Changes

1. Coverage under the Privacy Act and E-Government Act

While some problems with the Privacy Act became clear soon after the law was enacted, others have emerged over time. Most of these issues have arisen from a combination of the impact of new technology and the lack of updated guidance for agencies. These issues include:

1. Difficulty Categorizing Systems of Records

A major concern with the Privacy Act today centers on its most important term, "system of records," which is ill-suited to the current data

[52] In fact, the Committee on Authentication Technologies and Their Privacy Implications of the Computer Science and Telecommunications Board, Division at the National Research Council of the National Academy of Science titled a report "IDs—Not That Easy: Questions About Nationwide Identity Systems"

[53] http://www.cdt.org/privacy/authentication/030513interim.shtml

[54] OMB Memo m04-04 -- http://csrc.nist.gov/drivers/documents/m04-04.pdf

[55] Committee on Authentication Technologies and Their Privacy Implications, Computer Science and Telecommunications Board, Division at the National Research Council, The National Academies "Who Goes There?: Authentication Through the Lens of Privacy," National Research Council, 2003, p. 138.

environment. The definition of "system of records" excludes from the coverage of the Privacy Act information that is not regularly "retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual."[56]

This "system of records" definition is overly restrictive. As the PPSC suggested 30 years ago, the system of records requirement acts as an "on/off" switch for the Privacy Act's other requirements. Information that falls outside of the definition is not covered, no matter how it is used or misused.

Also, as discussed in detail above, the move to relational databases and distributed databases has changed how data is used and has clouded what is a "system of records" today.

For example, DHS did not original consider its ADVISE data mining program covered under the Privacy Act because ADVISE was not a system of records. The systems that ADVISE linked to were covered by the Act, but the narrowness of the concept of a "system of records" gave an incomplete picture of the privacy risks of ADVISE, which pulled data from several sources. (Because of scrutiny, DHS eventually suspended the system.[57]) The Privacy Act was certainly intended to address the full range of issues posed by a data mining programs like ADVISE, but the fact that DHS could claim the Act did not apply, illustrates how changes in technology have blurred the scope of the Act's most basic definition.

In its recent report, GAO directly cited as a major weakness in the law the Privacy Act's definition of "system of records" as not being universally applicable to the types of personally identifiable information collected by the government.[58] Because the definition is core to the main Act protecting personal information in the government, the weaknesses of what is not covered as a "system of record" calls the federal privacy protection framework as a whole into question.

In addition, there are growing concerns over the importance of non-personally identifiable data. As evidenced by recent statements from FTC and EU, the distinction between PII and non-PII is indeed "becoming less and less meaningful" in many contexts. Effectively addressing the issues

---

[56] 5 U.S.C. § 552a(a)(5).
[57] Ryan Singel, "DHS Data Mining Program Suspended After Evading Privacy Review, Audit Finds," Wired Threat Level Blog, August 20, 2007 http://blog.wired.com/27bstroke6/2007/08/dhs-data-mining.html.
[58] GAO, Privacy: Congress Should Consider Alternatives for Strengthening Protection of Personally Identifiable Information , GAO-08-795T (Washington, D.C.: June 18, 2008).

posed by the "system of records" definition will also require addressing the balancing act required to appropriately cover non-PII.

2. Weakness of Notice

a. Lack of Consistency in SORNs and PIAs

Concerns over the public's ability to understand and use government privacy notices dates back at least as far as the Privacy Protection Study Commission Report in 1977, which suggested that agencies were merely attempting to meet the letter of the law rather than its goals. The problem has worsened over time. In 1987, GAO found that more than half of the SORNs were inaccurate.[59] In 1990, a more comprehensive GAO study suggested that only 65% of systems covered by the Privacy Act had proper notice procedures.[60]

The requirement in the E-Government Act of 2002 to conduct Privacy Impact Assessments was intended to make available a lot more information about government data systems. However, implementation of PIAs has been uneven at best. The leader in implementing the PIA requirement, the Department of Homeland Security, has issued detailed rules about how PIAs should be completed and has regularly issued informative information about complex systems.[61] Meanwhile, the State Department issued a one and a half page PIA for the E-Passport System, a controversial system using RFID and biometrics technology.[62]

Privacy Act notices are intended to inform the public about privacy protections, but are fragmented and difficult to understand. In a 2008 report, GAO suggested that SORNs alone were not adequate to help even educated individuals understand how their privacy could be affected.[63] SORNs are published in the Federal Register, but are difficult to understand, overly vague and general, and reach only a very narrow audience. To be effective, notice must be relevant, easy-to-read and consistent. Current government privacy notices may serve some transparency purposes, but do little to explain information collection and use to the public.

---

[59] GAO, Privacy Act System Notices," November 30 1987, GAO/GGD-88-15BR http://archive.gao.gov/d29t5/134673.pdf,

[60] GAO, "Computers and Privacy: How the Government Obtains, Verifies, Uses and Protects Personal Data," August 1990, GAO/IMTEC-90-70BR

[61] http://www.dhs.gov/xinfoshare/publications/gc_1209396374339.shtm

[62] See http://www.cdt.org/security/identity/20070502rice.pdf

[63] GAO, Privacy: Congress Should Consider Alternatives for Strengthening Protection of Personally Identifiable Information , GAO-08-795T (Washington, D.C.: June 18, 2008).

c. Confusion over, and expansion of, the routine use exception

The issue that has caused the most concern over the history of the Privacy Act has been the frequent, seemingly standardless invocation of the "routine use" exception to override the Act's limits on reuse and sharing of information between agencies. The "routine use" exception was designed to allow agencies to share information in limited circumstances based on the frequency and administrative burden of obtaining individual consent. The Privacy Protection Study Commission raised major concerns about how the "routine use" exception was already being exploited to justify practices that went beyond the original intention of the Act. Successive administrations have become ever more accepting of this exception.

As technology has made information easier to share, the "routine uses" have proliferated and become less easy to understand in the SORNs. Routine uses are now so widely claimed and utterly unchecked that almost every Privacy Act Notice required by the law lists numerous routine uses, including vague boilerplate language confusing both citizens who want to understand what is happening to their data and the agency personnel responsible for it. For example, the Department of Defense regularly lists over 20 routine uses and then includes a Web link to a set of 16 "Blanket Routine Uses" that are included with every Privacy Act Notice it publishes.[64] While Congress left some of its intent on routine use open for interpretation, the current routine use system simply serves as a large loophole for the rest of the Act.

B. Lack of Leadership On Privacy

While some agencies have created strong and enduring privacy programs despite the changes in technology, many privacy failures have occurred due to the lack leadership of those individual federal agencies that have simply not devoted adequate attention to information privacy and security.. In June 2003, GAO issued a report entitled "Privacy Act: OMB Leadership Needed to Improve Agency Compliance." In that report, the GAO Identified deficiencies in compliance and concluded: "If these implementation issues and the overall uneven compliance are not addressed, the government will not be able to provide the public with sufficient assurance that all legislated individual privacy rights are adequately protected." Yet, criticism for failing to provide adequate oversight and guidance to agencies is not new. In 1983, the House Committee on Government Operations raised concerns that OMB had not updated its guidance in the first nine years of the Act's passage. The lack of guidance from OMB on Privacy Act implementation is more evident now precisely because of the increased challenges that agencies face in protecting privacy due to the rapid increase of technological change.

[64] The "Blanket Routine Uses" are available at http://www.defenselink.mil/privacy/dod_blanket_uses.html

ISPAB has noticed that the agencies that have done a better job of addressing privacy issues despite the ongoing lack of guidance and support are those that have created a privacy program with a strong Chief Privacy Officer at the helm. Some of these CPOs report to General Counsel, some to the Chief Information Officer and some even report directly to the Secretary of the Agency. Our discussions with these leaders have led us to believe that their success seems to be based on being organized in the right place in that agency and the resources afforded to them, than on a consistently repeatable formula across the government.

C. Other Issues Related To Privacy Policy

1. Ability to address new changes in technology as they are seen to greatly impact policy

Advancing technologies are quickly outpacing the Privacy Act's ability to keep up. The fact that it did a good job for 35 years is a testament to the fact that it was not written to fit the technologies of the day, but it is quickly becoming clear that technologies of today- data mining, location data, and sophisticated tracking technologies- simply do not fit within the definitions that were coined in 1974 and for which guidance was written in 1975. Agencies will need more regular guidance to be able to keep up with new technologies.

2. Expansive Nature of Potential Data Breaches and Impact on Individuals

The large increase[65] in high-profile data security breaches within the federal government in recent years makes it clear that the government is at risk of losing personal data on a massive scale. One of the most prominent federal data breaches occurred in 2006, when a laptop was stolen from a Department of Veteran's Affairs employee's home. Personal information like names, social security numbers, and addresses, were on the laptop and unencrypted. Despite policies stating that bringing this information home was a violation of policy, information security was made a priority at the VA and across the federal government. GAO released a report recommending that OMB create guidance that clearly determines what kinds of notice individuals affected by a government data breach should receive, and what kinds of services the government should offer them in order to minimize the risk of identity theft.[66] OMB has started down this path with a number of memos and other guidance to agencies on breach

[66] GAO, "Privacy: Lessons Learned about Data Breach Notification", April 2007 GAO-07-657

notification and other issues,[67] but how evenly agencies are implementing this guidance is still uncertain.

The more data is collected and the longer it is electronically stored, the greater the risk that it will be stolen, lost, or otherwise unintentionally disclosed. However, the Privacy Act's enforcement mechanisms generally only apply to intentional disclosures, and the Act offers little in the way of protections or remedies for unintentional loses, including those that may have been facilitated by poor data security practices.

V. Recommendations

A. Amendments to the Privacy Act of 1974 and Section 208 of the E-Government Act of 2002 are urgently needed.

1. Government privacy notices must be improved.

Current practices are not providing the level of transparency or clarity about major information systems that was expected by Congress when the Privacy Act and the E-Government Act were passed. ISPAB recommends the following solutions:

a. Best Practices for SORNs and PIAs are necessary.

Few people today read SORNs and PIAs, but those who do (Congress, GAO, IGs, journalists, advocates, other government agencies and their advisors and, in very rare cases, the courts) typically provide the necessary accountability for Privacy Act oversight. Today, SORNs are difficult to understand, and PIAs are entirely different from agency to agency. OMB should take advantage of best practices already in place at agencies such as DHS, USPS and the FTC, and publish best practice SORNs and PIAs and procedures to create and issue them.

b. Layered notices and Standardization of notices will create better oversight and accountability as the number and complexity of systems grow.

Industry has been experimenting with so-called "layered notices" that provide basic information to someone reading a privacy policy that then is linked to more detailed information. This snapshot allows reviewers to compare between policies more easily and creates a policy that is generally more readable to the general public. Agencies should be creating layered notices

[67] OMB's guidance can be found in the "Computer Security" and "Privacy" areas of guidance on the OIRA Web site — http://www.whitehouse.gov/omb/inforeg_infopoltech/#cs

and OMB should be including best practices of this kind in its guidance to the agencies.

Because SORNs and PIAs are difficult to read and find today, some experts have suggested urging a standardized structure so that it could then be more easily compared. In particular, the ability to sort and use by software agents to make sense of and compare the information provided would offer particular benefit in an era where these documents are released at a rate that it is difficult for interested parties to follow. ISPAB believes that standardization would improve accountability and should also be included in the best practices document.

c. A Privacy.gov site will also aid in oversight and accountability.

Finding a specific SORN or PIA can be difficult when a system of records is utilized by multiple agencies or is or has been known by several different names. GAO has suggested that OMB create a centralized Privacy.gov site to help solve this problem. ISPAB agrees. SORNs and PIAs could be made more easily locatable and searchable and privacy guidance could be posted there as well.

d. Breaking the routine use exception into its basic parts will help to ensure transparency

The confusing nature of the routine use has created uncertainty about what can and cannot be shared among agencies. In order to provide greater clarity, OMB could more clearly require agencies to state the authorization for the sharing of any record. If the authorization is tied to the creation of the system, this use should be consistent with the concept of "principal purpose(s)" in the Act. Agencies should still be required to disclose sharing outside the agency for principal purposes. If the sharing were not done for a principal purpose, but for some otherwise authorized purpose, then it should not be considered a routine use of data and it should be disclosed along with its authorization in statute or executive order. OMB could also mandate that secondary purposes internal to the agency be specified. This is not required today, but it would ensure that information is used only for specified purposes as required by the law.

While these changes would be an improvement, ISPAB would recommend eliminating "routine use" and just break it into its components thus ending the convoluted structure created in 1974. This would mean removing the concept of "principal purposes" and "routine uses" and replacing them with the following defined terms:

- The term "Primary uses" means a use of a record that is:
  A) necessary and relevant for a program or function for which the record was originally collected and
  B) authorized under either legislation or Executive Order of the President

- The term "Secondary uses" means a use of a record that is:
  A) necessary and relevant for a program or function other than that for which the record was originally collected and
  B) authorized under either legislation or Executive Order of the President.
- The term "Internal Sharing" means the sharing of a record within the government entity that created the record.
- The term "External Sharing" means the sharing of a record an entity other than the entity that created the record.

If this language were adopted, all sharing of information would be primary internal, primary external, secondary internal or secondary external, and rules would have to be developed to address sharing under each category. It would be clear that all uses aside from those exempted would need to be explicitly disclosed and tied to authorizing language.

2. Update the definition of System of Records to cover relational and distributed systems based on government use of, not holding, records.

Updating the concept of "system of records" is crucial for modernizing the Privacy Act. Some have suggested that all information that could be used to identify an individual at some time be covered by the Act. Yet this change would begin to cover almost every list of information in the government including email address books and even many word-processed documents. Another idea is to cover all personally identifiable information (PII) in all contexts, but even this then depends on an appropriate definition of PII. In fact, all solutions are a compromise of one kind or another in that agencies collect and use information in so many different ways, that it is difficult to develop a definition that is appropriate for all.

Of the possible compromises, ISPAB recommends a definition that would cover more systems than the current definition, but would provide agencies and OMB the opportunity to create use cases to ensure it is not overly broad.

3. Commercial Data Sources should be clearly covered under both the Privacy Act and the E-Government Act.

a. Direct Use of Private Sector Data without Merging should be considered a System of Records

Commercial information can and should play a key role in important government functions, including law enforcement and national security investigations. However, agencies relying on that data should have clear guidelines for its use—guidelines that both protect individual rights and ensure the information is reliable for the government purpose for which it is proposed to be used. Considering the harms that can occur when the

government makes decisions about individuals based on inaccurate or irrelevant data, it is imperative that the federal government develop better and more consistent rules for use of commercial data, regardless of whether the data is stored on government computers or stored on commercial systems. ISPAB agrees with GAO that the use of such systems should be covered under the Privacy Act.

b. Require Privacy Impact Assessments before time of contract for certain ongoing commercial data contracts

PIAs offer detailed information that can help policymakers understand privacy issues before procuring and finalizing information systems. Agencies should be required to undertake PIAs when contracting for most commercial data systems whether or not the information will be integrated into a government system of records. As more and more data services are made available online for subscription, it is important that privacy issues in using these systems are addressed so that policy can be written and employees can be properly trained on their use.

B. Government leadership on privacy must be improved.

1. OMB should hire a full-time Chief Privacy Officer with resources.

A Chief Privacy Officer is needed at OMB, with access to basic resources to provide government-wide privacy policy direction to the federal government. This OMB Privacy Office would oversee the privacy leadership within agencies and develop guidance and best practices.

2. Privacy Act Guidance from OMB must be regularly updated.

The fact that OMB has not reissued full Privacy Act guidance to the agencies since six month after the passage of the Act has led to an untenable situation. Agencies have been implementing the Privacy Act unevenly for years and the explosion of the use of new technologies with greater tracking capability has led to confusion among agencies. It should be a priority for the new CPO at OMB to issue new detailed Privacy Act guidance within six months and update it regularly every seven years after.

3. Chief Privacy Officers should be hired at all "CFO agencies."

The lack of privacy leadership is also apparent inside the agencies. Laws and policy creating senior privacy leaders at agencies have had been unevenly implemented. It should be made more clear that each agency should be required to hire a senior level Chief Privacy Officer, who will report to a senior agency official, such as the Secretary, the General Counsel or the CIO. This would reflect current business-sector practices and management approaches. Practically, this should be implemented at all of the so-called

"CFO agencies."[68] The OMB Privacy Office should oversee privacy protection implementation in agencies. In some cases, the CPO duties could be assigned to someone with other job roles- for example, in agencies where the only major source of PII is employee data.

4. A Chief Privacy Officers' Council should be developed.

The ISPAB has heard from a number of dedicated professionals serving the American people as CPOs at federal agencies. These officials generally seem to share the view that the CIO Council's creation of a subcommittee for privacy has been a positive step that allows them to meet and discuss best practices. However, they also have recommended that a separate Chief Privacy Officers' Council would provide a better structure for them to meet so that they can interact with each other and with other officials without depending on the CIO Council. ISPAB agrees that the public positioning of this group is important and will help raise the profile of privacy issues within the government.

C. Other changes in privacy policy are necessary

In issuing the above recommendations and options, ISPAB is not suggesting that these changes alone will address all future privacy issues as they arise. We do hope that this framework will make it easier for the federal government to respond to future privacy concerns quicker and more efficiently. We believe that this can begin in the following areas:

1. OMB should update the federal government's cookie policy.

OMB's current policies on cookies depend on bureaucratic speed bumps to protect user privacy. While this strategy has worked to some degree in the past, the utility of mechanisms such as cookies in Web 2.0 services will likely create greater incentive to circumvent user protections. Instead of banning the use of cookies, the government should be requiring clear opt-in consent mechanisms for the use of cookies. It should be the user's decision whether a cookie is set or not from a government Web site. The new policy need not be so prescriptive that it requires constant definition of cookies and prompting of users, but a clear consent to the storage and use of information to help provide a particular service, such as the "remember me" check boxes common on many commercial Web sites.

[68] In 1990, the Chief Financial Officers Act incorporated elements of business-sector best practices into the agencies, requiring all major agencies to create a chief financial officer (CFO) position. Under the CFO Act, OMB holds responsibility for financial management and improvement; a similar structure for Chief Privacy Officers should be implemented.

2. OMB should issue privacy guidance on agency use of location information.

One type of information that is particularly difficult to cover in any useful definition of system of records is location data. As the collection of location information increases, the policy and procedures for using location data will become more critical. OMB should be required to create guidelines for non-law enforcement use of location data by federal agencies.

3. OMB should work with US-CERT to create interagency information on data loss across the government

In our discussions with agencies on data loss incidents, it became clear to ISPAB that security and privacy personnel need more information from US-CERT about the incidents that other agencies report. Agencies are contributing information and could learn a great deal from their about the types of incidents to look out for; the quality of their own reporting; and other best practices. One means to help share this information among agencies would be to create a closed system to share information about data loss incidents.

4. Public reporting on use of Social Security Numbers

In its 2007 guidance to federal agencies, OMB rightly noted that an important step in preventing costly data breaches is "reducing the volume of collected and retained information to the minimum necessary."[69] OMB reflected SSN reduction policy in its update of E.O 9397, by striking "shall" and inserting "may," which many agencies had cited as their agency authority to collect and use SSNs. Another step that OMB required was an accounting of all collections of Social Security Numbers with a goal to minimize collections and retention when that information was not necessary to the purpose of the collection. One means to ensure that this policy is used to hold agencies accountable would be to publicly publish the number of SSN collections at each agency on a yearly basis. This could help create incentives and accountability by shining a spotlight on which agencies had failed to limit their SSN use and which had minimized SSN use.

[69] Clay Johnson III, "Safeguarding Against and Responding to the Breach of Personally Identifiable Information, OMB Memo M-07-16, May 22, 2007.