Testimony before the Senate Committee on Homeland Security and Governmental Affairs, Subcommittee on Oversight of Government Management, the Federal Workforce, and the District of Columbia on the "State of Federal Privacy and Data Security Law: Lagging Behind the Times?"

Peter Swire
C. William O'Neill Professor of Law, Moritz College of Law, Ohio State University
Senior Fellow, Center for American Progress Action Fund

July 31, 2012

Introduction

Chairman Akaka, Ranking Member Johnson, and distinguished members of the committee, thank you for inviting me to testify on "State of Federal Privacy and Data Security Law: Lagging Behind the Times?"

I am the C. William O'Neill professor of law at the Moritz College of Law at Ohio State University. In 1999 I was named chief counselor for privacy in the U.S. Office of Management and Budget. In that role, I was the first—and thus far the only—person to have governmentwide responsibility for privacy policy. As chief counselor for privacy, I worked extensively with the Privacy Act of 1974, helped institutionalize the practice of Privacy Impact Assessments for federal systems, and addressed many other privacy and cybersecurity issues affecting federal agencies. Since then I have continued to write and speak extensively on privacy and security issues.

For this testimony, committee staff requested that I address a range of issues concerning federal agency privacy and data practices. As the other testimony for this hearing demonstrates, there are many different privacy-related challenges facing federal agencies today. My testimony addresses four topics, with the key points set forth in the introduction:

1) **The Senate should promptly confirm the five nominees for the Privacy and Civil Liberties Oversight Board.** The most important short-term action the Senate can take on privacy is to confirm the five nominees for the Privacy and Civil Liberties Oversight Board, as voted out of the Judiciary Committee. All five nominees are supported by 9/11 Commission Co-Chairs Tom Kean and Lee Hamilton. Although there were dissenting votes in committee concerning the proposed Chairman David Medine, he is an outstanding and experienced nominee. By statute, only the chairman can hire staff, and the Senate should act promptly to put the Board into operation.
2) **Congress should create a federal chief privacy officer by statute to improve coordination of privacy policy across federal agencies.** A federal chief privacy officer would notably improve the clearance process within the executive branch for privacy policy, as well as help coordinate the many transborder privacy issues that arise in our world of pervasively global data flows. Without statutory support, existing agencies may stymie creation of that position. I suggest that the federal chief privacy officer might take the lead for nonclassified federal information technology systems, while the Privacy and Civil Liberties Oversight Board could take the lead for classified systems.
3) **There is an important loophole in the Privacy Act, but the problem can best be addressed by changes to the E-Government Act.** The proposed S. 1732 to update the Privacy Act correctly recognizes that the definition of "system of records" has an important loophole. The current definition applies only to records retrieved by name, and modern search engines often identify records even when the name does not appear in the search term. The proposed amendment would close the loophole but have the effect of requiring a far larger number of systems-of-records notices by federal agencies. In my view, this increase would create compliance burdens but not lead to significant privacy improvements. I believe a more promising approach would be to improve Privacy Impact Assessments under the E-Government Act of 2002. For instance, the Office of Management and Budget or an interagency council should post agency Privacy Impact Assessments to a unified web site so the public can compare agency assessments. Agencies should likely have a mechanism where public comments would be posted for the assessments. In addition, agencies could be required to respond to these public comments.
4) **The oversight process should focus more attention on the line between identified and deidentified data in federal agencies.** Specifically, the Federal Trade Commission has proposed a promising approach for defining deidentified data when held in the private sector. An important question is how that approach might be modified for use in federal agencies.

In summary, this committee is performing an important service by focusing attention on the privacy practices of federal agencies. I hope that the comments here will be of use to the committee in its oversight and legislative efforts.

The Senate promptly should confirm the five nominees for the Privacy and Civil Liberties Oversight Board

Before turning to the long-term issues of privacy and the federal government, there is one pressing privacy item for action by the Senate as soon as possible. The Senate should confirm the five nominees for the Privacy and Civil Liberties Oversight Board, as voted out of the Judiciary Committee. This past week's Senate vote on the cybersecurity bill makes confirmation even more urgent.

Currently, the Privacy and Civil Liberties Oversight Board is not in operation. The 9/11 Commission recommended implementing this type of board to increase oversight of the expanded information sharing practices among agencies adopted after the attacks of September 11, 2001. The Senate confirmed members of the Privacy and Civil Liberties Oversight Board in 2006, and the board began operation. Controversy emerged about the original board's lack of independence. As a result, a revised structure for the board was established in 2007 as part of the Implementing Recommendations of the 9/11 Commission Act. The revised structure creates staggered six-year terms for each of the five members and requires the chairman to work full time for the board.

No members of the board have been confirmed since that time. The Senate Judiciary Committee voted and approved all five nominees this May, but no date has been scheduled for floor action. Having a functioning Privacy and Civil Liberties Oversight Board is important under any circumstances to ensure regular and effective examination of the information sharing and privacy practices for homeland security and other antiterrorism activities.

The importance of implementing the board becomes even greater, however, due to the expanded information sharing in the proposed cybersecurity legislation. A key purpose behind that legislation is to enhance information sharing as a tool for fighting cyber attacks. A key safeguard is for the board to scrutinize this type of information sharing. In my view, putting the board in place should be a required component of approving cybersecurity legislation.

The full slate of nominees has received a strong letter of support from the Bipartisan Policy Center, signed by Tom Kean, former Republican governor of New Jersey, and Lee Hamilton, former Democratic congressman from Indiana.[1] Gov. Kean and Rep. Hamilton co-chaired the 9/11 Commission. In their letter this June, the authors wrote, "The Board is designed to play a crucial oversight role in preventing the intentional or accidental misuse of personal information across the government, and its establishment should be a high priority." They thus wrote to "advocate for the confirmation of the five nominees" to the board, all of whom have been reported out of committee.

I would also like to comment specifically in support of the nomination of David Medine to serve as the chairman of the board. Medine received dissenting votes on his nomination in committee, although there are no public reports of any basis for opposition or concern. I have known him professionally for more than 15 years. From 1992 to 2000 Medine was the senior civil servant expert on privacy at the Federal Trade Commission, serving as the associate director for financial practices. Shortly after, he became a partner at the leading law firm WilmerHale, where he worked with private-sector clients primarily on privacy and data security. In this position, he counseled clients on how to comply with complex privacy requirements. I believe this real-world compliance experience is highly relevant to realistic privacy protection. Medine has experience both in enforcing to protect privacy and in the burdens that exist when privacy rules are overly strict or badly drafted. This balanced experience makes him an outstanding person to chair the Privacy and Civil Liberties Oversight Board.

The statute creating the board requires the chairman to work full time. In addition, the statute allows only the chairman to hire staff, specifically that, "The chairman of the Board … shall appoint and fix the compensation of a full-time executive director and such other personnel as may be necessary to enable the Board to carry out its functions." Clearly, the board cannot carry out its work as the statute intends if there is no chairman in place. The Senate should act promptly to confirm all five nominees.

The importance of coordinating federal privacy policy

The committee asked me to write about my experience as chief counselor for privacy, including the merits of having a federal chief privacy officer to coordinate and oversee privacy policy across the federal government. I support the proposal by Sen. Akaka in S. 1732 to create such an office. The discussion here explains some key reasons that I support creating such a position. It then suggests how to structure such an office, with the federal chief privacy officer taking the lead on nonclassified federal information systems and the Privacy and Civil Liberties Oversight Board taking the lead on classified systems.

Why the federal government should have a privacy policy office

In a piece prepared for publication in the *Stanford Law Review* in 2000 (but which was not ultimately published), I explained the role that the chief counselor for privacy played during the intense privacy policy debates of the late 1990s.[2] Earlier this year I returned to the subject in a law review article on why the federal government should have a privacy policy office.[3] That article highlights the role such a privacy policy office would play in the interagency clearance process and in coordinating a unified approach to the large number of international privacy issues.

First, the chief privacy officer is important for the clearance process.[4] To ensure a unified administration position for congressional testimony, executive orders, and many other documents, drafts are circulated for clearance among the various agencies and components of the Executive Office of the President. Once comments are received, discussions are sometimes needed to resolve differences of opinion, with appeal to more senior officials if differences are not resolved at lower levels. In addition to these structured clearance procedures, agency experts on an issue such as privacy often get engaged earlier in the policy planning process in a variety of working groups and less formal methods of sharing expertise and views.

From my time as chief counselor for privacy, the number of privacy issues addressed by federal agencies is far greater than many people realize. Here is a list of the sorts of privacy issues that can arise in each of the cabinet departments:

- Department of Agriculture: Migrant worker records
- Departments of Defense and Veterans Affairs: Records of service members
- Department of Education: Education records, including for for-profit institutions
- Department of Energy: Smart grid
- Department of Health and Human Services: Medical records and many forms of human services records
- Department of Homeland Security: Numerous issues, including transportation safety and immigration
- Department of Housing and Urban Development: Public housing records
- Department of the Interior: National park reservations and other services provided online
- Department of Justice: Numerous issues
- Department of Labor: Records of union membership
- Department of State: International privacy issues
- Department of Transportation: Drone surveillance
- Department of Treasury: Financial privacy and money laundering

This list shows a wide variety of privacy issues and also that privacy issues emerge for new agencies over time. Surveillance by drones, for example, is becoming an important privacy issue as the Federal Aviation Administration permits expanded use of drones within the borders of the United States. For these kinds of emerging issues, I believe the expertise developed by a federal chief privacy officer would be quite useful.

Second, along with clearance, the executive branch needs effective coordination to develop and announce the administration position in international settings. Data flows today are pervasively global. We are reminded of this reality by the ongoing debates about the European Union's draft regulation on data protection. A very wide range of Internet and other private-sector data practices would be affected if that

regulation were to go into effect as currently written. For the public sector, there are also many cross-border issues such as passenger name records, law enforcement investigations, and many others. One of my current research projects analyzes how cloud computing, together with the widespread current adoption of encryption, is making international cooperation on law enforcement investigations much more important than it was in the past.[5] For the federal government, the increasing number and complexity of transborder privacy issues means that coordination of privacy policy would be very helpful.

From my time at the Office of Management and Budget and in the National Economic Council, there certainly are existing mechanisms for policy coordination. The National Economic Council and National Security Council are experienced at bringing together the relevant agencies to coordinate on complex policy problems. I believe these policy mechanisms, however, are not a good match for the ongoing privacy challenges. Resolving privacy issues often requires crosscutting expertise, drawing on specific domains, including information technology, law, business practices, and policy. When this complexity is added to the complicated interagency and international dimensions of the issue, the policy councils do not have the staffing and infrastructure to do a good enough job on managing privacy issues over time.

How to structure federal privacy policy leadership

I believe that Congress should create, by legislation, the office of the federal chief privacy officer and similarly should require each major agency to have a chief privacy officer.

The administration's recent Green Paper and White Paper on commercial privacy protection suggest the role that legislation can play here. The Green Paper in 2011 contained the idea of having an office in the Department of Commerce to coordinate privacy policy for commercial actors.[6] That office was dropped from the 2012 White Paper.[7] My sense is that this shift reflects the institutional difficulties in establishing a new office unless there is congressional support. Existing offices are reluctant to cede their current roles and budget. Congress mandated creation of the office of the chief privacy officer when it created the Department of Homeland Security, and the chief privacy officer in that department has been effective at having institutional support, compared to other agencies.

Based on my experience, I believe that the Office of Management and Budget is an effective location for the federal chief privacy officer. This fits the management responsibilities of the Office of Management and Budget. In 1999 after a survey found that privacy policies were lacking on many federal agency websites, we were tasked with defining acceptable privacy policies and then making sure that agencies posted them. That experience taught me and my staff the challenges of complying with rules and public scrutiny. That kind of experience helps the chief privacy

officer be more realistic when developing policy that other organizations are expected to follow.

One topic that could benefit from further discussion is how to integrate a federal chief privacy officer with the Privacy and Civil Liberties Oversight Board. I suggest some ideas here, but other approaches are worth considering. One way to split responsibilities is for the federal chief privacy officer to coordinate policy and oversight for unclassified information technology systems, while the Privacy and Civil Liberties Oversight Board would take the lead on classified systems. This apportionment of responsibilities would parallel the existing different requirements for classified and unclassified systems generally. In terms of function, the federal chief privacy officer would take the lead on clearance and other issues of crossagency coordination. The Privacy and Civil Liberties Oversight Board is designed to be independent of the executive branch, and thus would not play that interagency coordination role. Instead, its principal responsibilities would include oversight and investigation of data used in connection with antiterrorism efforts.

There is an important loophole in the Privacy Act, but the problem can best be addressed by changes to the E-Government Act

I now turn to the topic of amending the Privacy Act of 1974 and related statutes that create the framework for privacy protection in federal agencies. Chairman Akaka has taken a leadership position in proposing ways to update the Privacy Act for our modern information environment—including in S. 1732, the Privacy Act Modernization for the Information Age Act of 2011. As just discussed, I support that bill's approach to reconfiguring the management and coordination of privacy actions of federal agencies. I believe that a somewhat different approach may be more constructive, however, when it comes to amendments to the core definitions in the Privacy Act.

This portion of the testimony first provides a brief background about the Privacy Act of 1974. It next analyzes the "retrieved by name" loophole that S. 1732 seeks to close, before explaining why amendments to the E-Government Act of 2002 may be a more effective way to protect privacy while managing compliance costs of federal agencies.

Background on the Privacy Act of 1974

The Privacy Act was passed at the end of 1974, the year that President Richard Nixon resigned from office. Along with the Freedom of Information Act, it was enacted to address a pattern of secret government surveillance of American citizens. The history of this surveillance has been told before, but it is useful to periodically remind ourselves about actions such as the years of wiretapping of Martin Luther King, Jr., the domestic intelligence files created by the FBI on hundreds of thousands

of Americans, and the use of IRS tax records against the president's political enemies list.[8] We should learn from this history so we do not repeat it.

The Privacy Act as enacted was based on a 1973 report from the Department of Health, Education, and Welfare, which proposed five principles for a Code of Fair Information Practices:

1. There must be no personal data record-keeping systems whose very existence is secret.
2. There must be a way for a person to find out what information about the person is in a record and how it is used.
3. There must be a way for a person to prevent information about the person that was obtained for one purpose from being used or made available for other purposes without the person's consent.
4. There must be a way for a person to correct or amend a record of identifiable information about the person.
5. Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for their intended use and must take precautions to prevent misuses of the data.

As enacted, the Privacy Act essentially codified these principles. Individuals start with a baseline right that their personal information can only be disclosed with their consent. An important aspect of the law was to publish systems-of-records notices in the Federal Register so that the general public could learn about the existence and nature of federal databases. These notices provide details such as categories of records maintained, ways for individuals to access their own records, and routine uses that permit additional disclosures by the agency without individual consent.[9]

During my time at the Office of Management and Budget, I was the official responsible for answering questions about interpreting the Privacy Act, working closely with the Department of Justice office that publishes collections of Privacy Act cases. Based on my experience, the Privacy Act today continues to play a vital role in structuring federal agencies' use of personal information. The privacy-related actions of federal agencies today are far better than they would be without the Privacy Act. Systems-of-records notices help agencies consider what uses of information are lawful and appropriate, especially where the notices are thoughtfully crafted and not boilerplate. In my experience, agency officers working with the Privacy Act thoughtfully apply the law's fair information practices to individual disputes and situations as they arise.

The "retrieved by name" loophole in the Privacy Act

The core definitions of the Privacy Act today are the same as when the law was enacted 38 years ago. Our information processing technology today, however, is comprehensively different than it was in 1974, and so the committee is justifiably

exploring whether key definitions should be updated. S. 1732 addresses the most glaring weakness in the existing definitions, which can be called the "retrieved by name" loophole. My view, however, is that there may be more effective ways to address that problem, notably through changes to the E-Government Act of 2002.

The definition of "system of records" is central to the Privacy Act because it is the main device for dividing what is covered by Privacy Act requirements and what is not. In any regulatory system, the definition of the scope of coverage is especially important—if something is outside the scope of a law, then agencies or other regulated entities do not have to worry about the other details of compliance.

Since 1974 the Privacy Act has defined "systems of records" to mean "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." For each system of records, the agency must publish a systems-of-records notice in the Federal Register.[10]

The main problem with the definition of systems of records is that it applies only when "information is retrieved by the name of the individual." This approach made sense in the days when records were kept primarily in a physical file drawer. If you wanted to access a record, you would thumb through the alphabetical list of file folders until you found the right person. This approach also made a certain amount of sense in the early world of mainframe computers. The IRS, for instance, would organize tax records by name or Social Security Number. The Privacy Act covers that type of highly structured systems of records because the records are retrieved by name or the person's identifying number.

This definition, however, fails to cover many other ways that agencies handle personal information today. The 1977 Privacy Protection Study Commission gave the example of a search by the Veterans Administration by psychiatric diagnosis. Because the search was by diagnosis and not by name, the Privacy Act simply did not apply to the search.[11] In essence, the Privacy Act definition applies to structured record sets listed by name but not to the other ways agencies can use records to identify and then act on individuals.

Due to increased speed and capacity of computer searches and data mining over the years, this gap in the Privacy Act's coverage has widened significantly. Because search is a daily part of our lives today, sometimes it is hard to remember that Google was not incorporated until 1998. Individuals and federal agencies today complete an enormous number of searches without use of a name, but people's names still pop up in the results. Data mining takes that concept even further—federal agencies sift through innumerable records in order to spot patterns and turn up suspects or individuals that are of interest for one reason or another. But the Privacy Act simply does not apply to the vast bulk of records where there is no organized retrieval by name or number.

To address this gap, S. 1732 would broaden the definition of systems of records to include "a group of any records maintained by, or otherwise under the control of any agency that is used for any authorized purpose by or on behalf of the agency." The proposed amendment recognizes how records are actually retrieved today, often without explicitly searching by name or identifying number. The proposed amendment would close the loophole that has been recognized since the 1970s.[12]

Under the new approach, the key trigger for Privacy Act coverage would be what qualifies as a record. The definition of "record" focuses on each individual, rather than on how records are grouped in an agency's filing system. Under the Privacy Act, the term "record" applies broadly to "any item, collection, or grouping of information about an individual that is maintained by an agency." The act provides examples of what count as records such as "his education, financial transactions, medical history, and criminal or employment history." Finally, a record "contains his name, or the identifying number, symbol, or other identifying particular assigned to the individual, such as a finger or voice print or a photograph."

The proposed amendment would close the "retrieved by name" loophole but would quite possibly also lead to an enormous increase in the number of systems-of-records notices. S. 1732 would apply to a group of any records under the control of an agency. My concern is that there would be too many groups of any records. Records today are gathered and used for many purposes. Under the proposed revisions to the Privacy Act, agencies would have to go through the bureaucratic requirements of systems-of-records notices for each of those groups. These notices provide important functions such as notifying the public and ensuring that the full set of Privacy Act fair information practices apply. The Information Security and Privacy Advisory Board's 2009 report on federal privacy protection, however, found that the notices "are difficult to understand, overly vague and general, and reach only a narrow audience."[13] I believe the Congress should consider other alternatives before acting to increase the number of systems-of-records notices in this way.

Consider improving Privacy Impact Assessments rather than directly amending the Privacy Act loophole

The discussion of the "retrieved by name" loophole shows an important flaw in the Privacy Act's goals of providing notice about agency privacy practices and ensuring consideration of privacy risks. Rather than amending the Privacy Act, however, I think that better progress can likely be made by improving the E-Government Act of 2002.

The E-Government Act requires agencies to issue Privacy Impact Assessments in connection with the "development or procurement of new information technology." Section 208 of the act requires Privacy Impact Assessments to be commensurate with the size of the information system, the sensitivity of the identifiable information, and the risk of harm from unauthorized release.

In considering the vast range of data used by federal agencies, my sense is that that the trigger for requiring a Privacy Impact Assessment is more practical than the proposed trigger for requiring a systems-of-records notice. A Privacy Impact Assessment is required when developing or procuring a new information technology system. In this way, the Privacy Impact Assessment is built into an ongoing process such as a procurement. Ideally, the assessment is completed early enough in the process to identify privacy risks, leading to a more effective and less privacy-intrusive system. In addition, the Office of Management and Budget has issued guidance under the E-Government Act that contains commonsense exceptions to the requirement that an agency do a Privacy Impact Assessment, including for minor changes to a system that do not create new privacy risks.[14]

By contrast, the proposed amendment would trigger a systems-of-record notice for a group of any records controlled by the agency. My concern is that the number of systems-of-record notices would need to climb substantially to cover this apparently very broad language. The Office of Management and Budget has authority under the E-Government Act to create pragmatic exceptions to when a Privacy Impact Assessment is required, but it is not clear to me that the Office of Management and Budget has similar authority under the Privacy Act. In addition, the Privacy Act does not have the risk-based approach of the E-Government Act, where the level of privacy work by the agency is supposed to be commensurate with the privacy risks.

My related concern is that increasing the number of systems-of-record notices would not actually improve privacy protection. At least ideally, the goal of a Privacy Impact Assessment is to do a nuanced examination of the privacy risks in a new procurement or computer system. This sort of nuanced examination, however, is unlikely to occur if an agency has to slog through a huge number of routine Privacy Act systems-of-record notices. If the number of notices climbs sharply, I fear that agencies will adopt too much of a "check the box" approach to privacy protection, simply filing Privacy Act notices that are uninformative and do not adequately address actual privacy risks.

In 2003 the Office of Management and Budget issued a Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002.[15] This guidance does a straightforward and reasonable job of implementing the E-Government Act as written. I have concerns, however, about how well the guidance has been implemented over time.

Going forward, this subcommittee and committee may find it useful to conduct oversight specifically on implementation for Privacy Impact Assessments of the E-Government Act and the Office of Management and Budget guidance. My sense of implementation of the assessments is similar to that found by the Information Security and Privacy Advisory Board. The Department of Homeland Security has done a notably good job in preparing and publishing Privacy Impact Assessments, in

no small part due to the visible leadership and responsibilities of the department's chief privacy officers, including Mary Ellen Callahan, who is testifying in this hearing today. Other agencies, however, have done a more superficial job in drafting their assessments. I am not aware of any major visible discussion about how to bring the quality of those other agencies up to the quality at the Department of Homeland Security.

I have two suggestions for improvement to the privacy parts of the E-Government Act. The first concerns making it easier to find and compare agency Privacy Impact Assessments. The act directs agencies to submit their assessments to the Office of Management and Budget. They are also directed to make their assessments publicly available, with certain exceptions for national security and other exceptions. Notably, these two requirements do not seem to be currently linked—I can find no easy way to find the Privacy Impact Assessments of different agencies in order to compare them. I think it would likely improve the quality and consistency of the assessments if the Office of Management and Budget or one of the interagency councils created a process for posting agency the assessments to a unified site that is publicly available.

Second, the E-Government Act could have more effective methods for public comment and input. As a first step, agencies should have a mechanism where public comments would get posted for PIAs. In addition, agencies could be required to respond to comments. The idea here is not to create full Administrative Procedure Act notice-and-comment, where a rulemaking cannot go forward until the comments are complete. Instead, my suggestion is a lighter-touch approach, where the agency would publish the public comments and give some response. This sort of nudge to an agency is consistent with the light-touch or nudge approach to regulation that Office of Information and Regulatory Affairs Director Cass Sunstein has brought to the Office of Management and Budget.

The oversight process should focus more attention on the line between identified and deidentified data in federal agencies

One increasingly important issue over time is determining how to draw the line between data that is identified or not. Privacy requirements apply where the links to a specific person are clear enough. By contrast, those requirements do not apply where the links are not clear enough such as where enough details are removed so that the information can be considered deidentified. The issue of deidentification has begun to receive significantly more attention in connection with personal privacy, as reflected this year in the administration's White Paper and the Federal Trade Commission's privacy report. My discussion here suggests that the oversight process should focus more attention on the line between identified and deidentified data in federal agencies. Specifically, the Federal Trade Commission has proposed a promising approach for defining deidentified data when held in the private sector.

An important question is how that approach might be modified for use in federal agencies.

This spring the administration released a White Paper called "A Framework for Protecting Privacy and Promoting Innovation." The White Paper applies to personal data held in the private sector. The title reflects the risks to individuals if privacy is not protected effectively. It also reflects the importance of creating good information rules in order to foster innovation and growth in our information economy.

The issue of deidentified data creates a vital opportunity to meet both goals—protecting privacy while using data for innovation, growth, and the other goals of the private and public sectors. At least in theory, deidentified data allows us to have our cake and eat it, too. With deidentified data, we strip out the name and other information that reveals identity, but we nonetheless can process the data, do research, discover patterns, and innovate in how we respond to the information.

In recent years, we have learned a great deal about when and how it is possible to "re-identify" data—to link a person's name with supposedly deidentified data. Two big trends have made it harder to keep information deidentified. First, searching on the Internet has gotten much better and more accurate. Today's search engines let anyone link together tidbits from previously hard-to-link data sources. Second, the amount of information on the web about a typical person has grown astronomically, including all of the personal details on a person's blog or Facebook page.

The combination of efficient search tools and lots of data means that there is a higher likelihood today that a person's records can be reidentified even if the name and other traditional identifiers are deleted. For instance, a deidentified medical record might state that a person in Ohio had minor hand surgery on April 3. In the past it would have been difficult or impossible for an outsider to figure out the patient's name. Today an online search might turn up a social network thread about the hand surgery—there are multiple such surgeries in Ohio each day but not that many. A bit of follow-up research using the rest of the supposedly deidentified information might easily pinpoint the person who had the surgery.

As experts have analyzed these facts about reidentification, some have concluded that the entire effort to deidentify data has failed because of the risk of linking information back to the individual.[16] Others have emphasized the limited actual success of reidentification efforts in practice and found that the benefits such as research and innovation are so great that they outweigh the privacy risks.[17]

In response to public comments on the issue of deidentification, the Federal Trade Commission in its privacy report this spring proposed a promising approach for treating data as deidentified. The FTC provides what amounts to a safe harbor where: "(1) a given data set is not reasonably identifiable; (2) the company publicly commits not to re-identify it, and (3) the company requires any downstream users

of the data to keep it in de-identified form." A key part of the approach is that the entity holding the data promises not to reidentify it. For instance, even if the entity could theoretically investigate who had the hand surgery on April 3, it won't do the investigation, and the data can be properly treated as deidentified.

I believe a similar approach could help federal agencies gain benefits from using data while holding it in deidentified form. The precise Federal Trade Commission approach will not work, however. Enforcement of that approach is based on the company's public commitment not to reidentify the data. A violation of that commitment is enforceable under Section 5 of the Federal Trade Commission Act, which prohibits unfair and deceptive trade practices. That act applies only to commercial actors, though, and not federal agencies.

The question is how to take this approach of promising not to reidentify and apply it to federal agencies. This is a novel question, and I do not know today how best to translate the Federal Trade Commission approach to federal agencies. I believe it is a worthwhile endeavor, however, because such an approach could open agencies to more of the modern benefits of using data, while also protecting privacy and reducing compliance costs with privacy requirements. Federal agencies also face the issue that information might be reidentified in some instances for law enforcement, national security, or related purposes. To address this possibility, one might require agencies to notify the Privacy and Civil Liberties Oversight Board (assuming it is up and running) if they reidentify data for national security or related reasons.

The ability to deidentify is becoming more technically challenging, while the need for effective deidentification is increasing. The Federal Trade Commission has proposed an approach that combines promises not to reidentify with the available technical measures. This fall I will be conducting a project on deidentification with the Future of Privacy Forum, seeking to identify and improve best practices in the area.[18] Along with efforts in the private sector, this committee in its oversight role can encourage the Office of Management and Budget and other federal agencies to create guidance and best practices for deidentification in the public sector.

Conclusion

I commend the committee for its attention to these important issues of privacy protection and federal agencies. Thank you for the opportunity to testify, and I welcome any questions you may have.

Biographical information

Peter Swire is the C. William O'Neill professor of law at the Moritz College of Law at Ohio State University. He is Senior Fellow with the Center for American Progress

Action Fund and the Future of Privacy Forum, and Policy Fellow with the Center for Democracy and Technology.

Swire began writing about privacy and the law of the Internet in the mid-1990s. In 1998 he was the lead author with Robert Litan of *None of Your Business: World Data Flows, Electronic Commerce, and the European Privacy Directive*, published by the Brookings Institution. In 1999 he was named chief counselor for privacy in the U.S. Office of Management and Budget. In that role, he was the first—and thus far the only—person to have governmentwide responsibility for privacy policy.

As chief counselor for privacy, Swire's activities including being the White House coordinator for the proposed and final HIPAA medical privacy rules and chairing a White House task force on how to update wiretap laws for the Internet age. He participated in the negotiation of the Safe Harbor agreement for transborder data flows between the European Union and the United States. For federal agencies, he oversaw the Office of Management and Budget's guidance on interpretation of the Privacy Act and chaired the privacy subcommittee for the federal Chief Information Officer Council.

Swire returned to teaching law after his work at the Office of Management and Budget and has continued to write and speak extensively on privacy and security issues, with publications and speeches available at www.peterswire.net. In 2009 and 2010 he was special assistant to the president for economic policy, serving in the National Economic Council under Lawrence Summers. In 2010 he again returned to teaching law at Ohio State University. He lives in the D.C. area.

Endnotes

[1] Letter to Senate Majority Leader Harry Reid and Senate Minority Leader Mitch McConnell from BPC Homeland Security Project Co-Chairs Lee Hamilton and Tom Kean, June 27, 2012, available at http://www.scribd.com/doc/98470241/Letter-to-Senate-Majority-Leader-Harry-Reid-and-Senate-Minority-Leader-Mitch-McConnell-from-BPC-Homeland-Security-Project-Co-Chairs-Lee-Hamilton-and-T.

[2] Peter P. Swire, "The Administration Response to the Challenges of Protecting Privacy" (Palo Alto, California: Stanford Law Review Symposium on Privacy, 2000), available at http://www.peterswire.net/pspublications-unpub.htm.

[3] Peter Swire, "Why the Federal Government Should Have a Privacy Policy Office," *Journal of Telecommunications & High Technology Law* 10 (41) (2012), available at http://ssrn.com/abstract=1960634.

[4] Swire, "The Administration Response to the Challenges of Protecting Privacy."
[5] Peter P. Swire, "From Real-Time Intercepts to Stored Records: Why Encryption Drives the Government to Seek Access to the Cloud," available at http://ssrn.com/abstract=2038871.
[6] U.S. Department of Commerce, *Commercial Data Privacy and Innovation in the Internet Economy: A Dynamic Policy Framework* (2010), available at http://www.commerce.gov/node/12471.
[7] The White House, *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy* (2012), available at http://www.whitehouse.gov/sites/default/files/privacy-final.pdf.
[8] For one discussion of the history, see Peter P. Swire, "The System of Foreign Intelligence Surveillance Law," *George Washington Law Review* 72 (1304) (2004), available at http://ssrn.com/abstract=586616.
[9] 5 U.S.C. §552a(e)(4).
[10] 5 U.S.C. §552a(e)(4).
[11] Privacy Protection Study Commission, "The Privacy Act of 1974: An Assessment" (1974), available at http://epic.org/privacy/ppsc1977report/appendix4.html. This document provides an example of a Veterans' Administration search by psychiatric diagnosis that was not covered by the Act. For in-depth discussion of the definitions that trigger privacy requirements, see Paul M. Schwartz & Daniel J. Solove, "The PII Problem: Privacy and a New Concept of Personally identifiable Information," *New York University Law Review* 86 (1814) (2011).
[12] The approach in S. 1732 is similar to the approach in the work of the Center for Democracy and Technology on how to update the Privacy Act. The Center for Democracy and Technology would delete the current definition of system of records, thus expanding the scope of the Privacy Act to the broader range of agency actions affecting records. See E-Privacy Act Amendments Wiki, available at eprivacyact.org (last accessed July 2012).
[13] Information Security and Privacy Advisory Board, "Toward a 21st Century Framework for Federal Government Privacy Policy" (2009), available at http://csrc.nist.gov/groups/SMA/ispab/documents/correspondence/ispab-report-may2009.pdf.
[14] Office of Management and Budget, *Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002* (The White House, 2003), available at http://www.whitehouse.gov/omb/memoranda_m03-22/.
[15] Ibid.
[16] Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization" *UCLA Law Review* 57 (1701) (2010), available at http://ssrn.com/abstract=1450006.
[17] Jane Yakowitz, "Tragedy of the Data Commons," *Harvard Journal of Law and Technology* 25 (2011), available at http://ssrn.com/abstract=1789749.
[18] The comments on deidentification here draw in part on material that I submitted to the Department of Commerce in its request for comments on the privacy multistakeholder process. See National Telecommunications and Information Administration, *Comments on Multistakeholder Process* (U.S. Department of

Commerce, 2012), available at http://www.ntia.doc.gov/federal-register-notice/2012/comments-multistakeholder-process; Peter Swire, "Keynote – Setting the State: How De-Identification Came into U.S. Law and
Why the Debate Matters Today," Speech at Future of Privacy Forum, Conference on De-Identification, Washington, D.C., December 5, 2011, available at http://www.peterswire.net/psspeeches2011.htm.