Applying Federal Data Security Policy to Statistical Agency Practices[1]

Shelly Wilkie Martinez and John W. Barkhamer
U.S. Office of Management and Budget, Washington, DC 20503

Abstract

Many in the media dubbed both 2005 and 2006 the "Year of the Data Breach," given well-publicized breaches across the private sector, academia and government. The Office of Management and Budget (OMB) led Federal efforts with the issuance of a key policy memorandum addressing data breach notification; unnecessary collection and retention of personally identifiable information, e.g., Social Security numbers; and a number of other privacy and security aspects. While government-wide in scope, the guidance understandably has a distinct effect on statistical agencies whose missions often involve collecting personally identifiable information (PII) and whose success rests on maintaining the public's trust. This paper focuses on how these OMB requirements affect statistical agency programs, particularly in tandem with statistical agency implementation of OMB guidance on the Confidential Information Protection and Statistical Efficiency Act of 2002.

Key Words: data breach, information policy, personally identifiable information, confidentiality

1. Statistical Agencies and Confidentiality – Historical Roots in the United States

Section 2 of Article 1 of the U.S. Constitution mandated a decennial census for the purposes of allocating seats in the House of Representatives. As statistical techniques advanced and the information needs of policymakers grew during the nineteenth century, the question arose of what measures the Federal Government should take to protect individuals' confidentiality. The longstanding question arose again in the last century as the Federal Government collected new kinds of information to develop and carry out a range of policies, such as controlling communicable diseases and administering an income tax system. In the last decade, the question of how best to protect the confidentiality of information held by the Federal Government continued, along with changes in technology that had broad effects.

2. Setting the Stage – The Early 2000s

2.1 Federal Public Sector Context and Existing Information Security Policy

During the information technology (IT) boom of the 1990s and up through today, the Federal Government has sought to match the public's increasing expectations for receiving enhanced services through information technology. A veteran expects a Department of Veterans Affairs (VA) doctor in Denver treating her to instantly retrieve records through the Internet on treatment she received at the VA medical center in Houston. A statistical agency survey participant expects a field representative who visits him to be carrying a laptop instead of a pen and paper and an Internet response option when receiving a mail survey. While the President's Management Agenda applied a unifying theme of advancing citizen-centric delivery of government services, Federal agencies applied technology to meet their diverse missions and in the context of their individual organizational structures.

As the Federal Government improved citizen service delivery through the use of the Internet, mobile electronic devices, and other information technology advances, the risks to information security and individual privacy was growing. The veteran could, among other concerns, face significant embarrassment and potentially other personal loss if she were on a list of drug abuse treatment participants accidentally posted to a public Web site. The survey participant could face serious economic loss if the Internet connection were not secure or if the laptop containing his personal information were lost or stolen, as well as other harms depending on the nature of the survey. Both the veteran and the survey

[1] This views expressed in this paper are those of the authors and do not necessarily represent those of OMB.

participant would expect reasonable, and perhaps extensive, safeguards to have been taken by the cognizant Federal agency and would expect notification of the losses and remediation.

As the many risks associated with technological advances have become evident, the statutory framework underlying Federal information security policy has been refined. Building off of the broad agency information security responsibilities in the Privacy Act of 1974, the Paperwork Reduction Act of 1995, and other legislation, the Federal Information Security Management Act of 2002 (FISMA) further detailed the information security statutory framework. The responsibilities included integrating information security into new IT investments, employee training, and periodic testing of controls.

As part of the emerging information security framework, the National Institute for Standards and Technology published guides for agencies on a range of issues, including user authentication, risk management, and configurations. The statutes assigned overall information security policy and oversight to the Director of the Office of Management and Budget (OMB). Agency heads and agency chief information officers (CIOs) bear responsibility for implementing requirements of those acts.

The risks associated with the loss or inappropriate disclosure of individual's personal information and the need to mitigate those risks were not new during the IT boom. The Privacy Act of 1974 requires agencies to "establish appropriate administrative, technical and physical safeguards to insure the security and confidentiality of records and to protect against any anticipated threats or hazards to their security or integrity which could result in substantial harm, embarrassment, inconvenience, or unfairness to any individual on whom information is maintained." Similarly, the Privacy Act assigns oversight responsibilities to the OMB Director. Agency statutes have also provided requirements to ensure the adequate protection of personal information, such as for the Internal Revenue Service and the Social Security Administration. The statutory and policy infrastructure emphasized consistent approaches to prevention and safeguarding, but as we will demonstrate, not to responding to information loss.

2.2 Private and Public Sector Experiences with Data Breaches

Just as the nation's privacy and security laws and policies have persisted over time, the public's concern to protect individual privacy is a long-standing part of American culture, reflected in the Privacy Act and other privacy statutes. Therefore, it was of growing public concern that by 2007, data breaches in the private and public sectors had become regular news stories, with the vulnerability of the Internet and mobile electronic devices forming a recurring theme. These data losses were reportedly suffered by both large and small private sector companies across business sectors, with legal and regulatory security and privacy requirements placed on the companies in many cases (e.g., healthcare under the Health Insurance Portability and Accountability Act as overseen by the Department of Health and Human Services and financial services under the Financial Modernization Act of 1999 as enforced by the Federal Trade Commission, seven other Federal agencies, and the states). In December 2006, *Federal Computer Week* declared 2006 the "Year of the Breach" due to the frequency and magnitude of Federal agency data breaches, which were mirrored in the private sector as well.[2]

Episodic stories of private sector data losses culminated in the spring and summer of 2006 in numerous press accounts reporting on the theft of a laptop containing 26.5 million individuals' personal information from the home of a VA employee and a string of relatively smaller, but still substantial, losses by other Federal agencies. In July 2006, the Inspector General's Office at the Department of Transportation suffered the theft of a laptop containing 133,000 individuals' personal information. In August 2006, a Department of Education contractor put the personal information of 21,000 student loan portal users at risk by accidentally enabling other users to view their accounts. [3]

Along with the security risks of the Internet and mobile electronic devices, agencies' responses to the breaches received closer attention. Agencies wrestled with how to report breaches up through their leadership, when to send out letters notifying individuals of breaches, what to include in such letters, what the risks of misuse of the data were, and when and what remediation to provide individuals.

[2] http://www.fcw.com/print/12_44/news/97098-1.html
[3] http://www.govexec.com/dailyfed/0806/082406p1.htm

3. Federal Data Breach Policy

3.1 The Rationale

As the public and Federal agencies observed a growth in the frequency and magnitude of loss of or unauthorized access to personal information, policymakers realized that they would need to quickly develop response policies and enhanced security measures to match the increased risks associated with technology advances under the statutory framework.

Building on the emerging recognition that identify theft can result in damage to victims, codified in the Identity Theft and Assumption Deterrence Act of 1998, the President's Identity Theft Task Force in 2007 noted that Federal agencies over the past year did not have "comprehensive formal guidance on how to respond to data breaches."[4] Based on the emerging consensus that the existing privacy and information security framework required a modest but important augmentation to adequately address this widely reported on problem, manifest in findings of the 17-agency task force and the lessons learned from VA and other agencies, OMB issued Memorandum 07-16 *Safeguarding Against and Responding to the Breach of Personally Identifiable Information* (Memorandum 07-16) in May 2007.[5] This memorandum provided an overarching framework for Federal agencies to respond to data breaches, balancing the need for guidance with the need for agencies to adapt the requirements to the specific circumstances each faced. Within the statutory foundation for privacy and securing Federal information, OMB provided clarifying guidance to address risks in the constantly changing Federal information environment.

3.2 The Requirements

In order to address these emerging risks, the memorandum required agencies to develop and implement a breach notification policy vis-à-vis individuals potentially affected and the public within 120 days. Each agency policy needed to cover the elements listed below, with the agency head ultimately making a risk-based decision when executing the policy.

- Whether breach notification is required;
- Timeliness of the notification;
- Source of the notification;
- Contents of the notification;
- Means of providing the notification; and
- Who receives notification.

Recognizing the potential "chilling effects" of notices and the costs to individuals and businesses of responding to notices, the memorandum also states agencies "should exercise care to evaluate the benefit of notifying the public of low impact incidents."

In addition to the external breach notification requirement, the new policy required agencies to report all incidents involving personally identifiable information (PII) within one hour of detection or discovery to the Department of Homeland Security U.S. Computer Emergency Readiness Team (US-CERT), which is charged with protecting the country's cyber infrastructure. If the breach is a result of an individual or organization intentionally attempting to gain access to Federal information, US-CERT would assist in coordinating a response.

The memorandum also requires agencies to complete a review of the use of Social Security numbers and other personally identifiable information and reduce their unnecessary collection and use. OMB, the Social Security Administration, and other agencies are exploring alternatives to agency use of Social Security numbers as unique identifiers in Federal programs. For Federal employees, OPM is leading the effort to develop policy for unique identifiers in Federal information systems and on related forms.

Memorandum 07-16 also restates previously existing information security guidance for sensitive Federal information, including sensitive personally identifiable information. The requirements include:

[4] http://www.idtheft.gov/reports/StrategicPlan.pdf
[5] http://www.whitehouse.gov/omb/memoranda/fy2007/m07-16.pdf

- Encryption on mobile computers and other electronic devices carrying sensitive data;
- Two-factor authentication (i.e., a typical user password requirement and a second, independent form of authentication, such as a token with access codes) for remote computer access;
- Time-out function for remote access requiring user re-authentication after thirty minutes of inactivity; and
- Logging of computer readable data extracts from databases holding sensitive personally identifiable information and verifying each extract has been erased or is still required within 90 days.

4. Statistical Agency Context

4.1 Data Protection in Statistical Agencies

Statistical agencies are inherently in the "information" business, so the OMB policy would naturally be expected to affect them more so than many other agencies that have other kinds of missions. At the same time, one also might assume that statistical agencies would be better positioned to implement such policy, given the centrality of data policies and procedures to their activities. A closer look reveals some important differences between the statistical agency "as is" and the new policy requirements, requiring sometimes significant effort to comply.

Some statistical agencies were "born" with confidentiality in their authorizing statutes (e.g., the Bureau of Transportation Statistics and the Bureau of Justice Statistics).[6] Others came to confidentiality statutes in their relative youth (e.g., the Census Bureau and the National Center for Health Statistics). Still others have had to rely on more general purpose statutes such as the Privacy Act or on internal policy and practice (e.g., the Bureau of Labor Statistics (BLS)). Some have had their statutes tested in court (e.g., the Census Bureau). But the value of being able to promise confidentiality to respondents is well established, so all of the principal statistical agencies have invested in efforts to ensure that confidentiality assurances are backed up by practice.

As a tool for reassuring respondents, confidentiality works hand in glove with the principles of "statistical use only," and "functional separation." The former connotes the importance of individually identifiable data only as it pertains to the developing and reporting of aggregate or anonymous information not intended to be used, in whole or in part, for making a decision about an individual that is not an integral part of the particular statistical project. Functional separation refers to separating the use of information about an individual for a statistical purpose from its use in arriving at an administrative or other decision about that individual.[7] In determining privacy impacts, applying these concepts is often seen as a significant risk mitigator since the data are not meant to be used in ways that affect the individual.

The "parent" law under which FISMA was included – the E-Government Act of 2002 – also included the Confidential Information Protection and Statistical Efficiency Act (CIPSEA), which provides strong confidentiality protections for information collected for exclusively statistical purposes under a pledge of confidentiality. This law provides a uniformly high floor of protection for statistical data and directly addresses the problem that some statistical agencies, such as BLS, faced without a confidentiality statute of their own. Commensurate with the strong statutory protections and penalties that the law provides, agencies are required to institute appropriate safeguards and procedures for protecting information collected under CIPSEA, as specified in guidance issued by OMB in 2007.[8]

So in the world of 2005-2006, statistical agencies were either operating under their own confidentiality regimes or working to incorporate CIPSEA. Many had formal data protection programs, including designated chief privacy or confidentiality officers, policies, procedures, training, and materials specially designed for respondents. Other authors

[6] For purposes of this paper, we limit discussions to those "principal" statistical agencies, i.e., agencies whose primary mission is statistical activities. These agencies include the Census Bureau, the Bureau of Labor Statistics, the National Center for Health Statistics, et al. We distinguish this subset from the over 80 Federal agencies with statistical activities sufficient to meet criteria for inclusion in OMB's "Annual Report to Congress on Statistical Programs of the United States Government" (http://www.whitehouse.gov/omb/inforeg/statpolicy.html#sp).

[7] *Personal Privacy in an Information Society: The Report of the Privacy Protection Study Commission* transmitted to President Jimmy Carter on July 12, 1977.

[8] http://www.whitehouse.gov/omb/fedreg/2007/061507_cipsea_guidance.pdf

have written about intentional efforts to instill a "culture of confidentiality" in these agencies.[9] Such efforts have included attention to the full data lifecycle, from data collection to product dissemination. For example, specialization in the area of disclosure avoidance in statistical products remains a focal point of a permanent interagency group, the Confidentiality and Data Access Committee, which was formed under the Federal Committee on Statistical Methodology.[10]

4.2 Statistical Agency Organizational Factors

4.2.1 Statistical Agency Business Practices

At the same time, statistical agencies were experiencing new challenges to data protection brought on by changes in business practices. Data collection has always been an inherently decentralized process, whether because of mail-out surveys requiring "paper" sent out and back via the mail or an in-person cadre of interviewers geographically dispersed. However, the IT boom led to greater use of computer-assisted personal interviewing, meaning that interviewers began to carry around one or more laptop computers as well as computer disks, flash drives, and paper, all containing potentially large amounts of personal information. Laptops are expensive electronic devices and thus an attractive target for thieves to take, as well as for separating interview staff who may wish to keep them. The small size of disks and flash drives makes them candidates for being lost, as well.

Technology also challenged built-in confidentiality protections in publicly available datasets, causing disclosure avoidance experts to add additional protections. In some cases, these meant the release of analytically less useful public micro-data sets. As a result of this and increased researcher technological capacity, agencies were under pressure to expand qualified researcher access to restricted datasets.

In addition, statistical agencies were frequently contracting for data collection and processing, as well as specialized services such as locating potential respondents. For a variety of reasons, including departmental contracting policies and specialization in the survey profession, many of these contractual relationships led to subcontracting relationships. Data from a single survey might be "touched" by four or five organizations in addition to the agency itself. Clearly, statistical agencies had their hands full ensuring data protection given their internal operating environments.

4.2.2 Statistical Agency Departmental Placement

Contracting was not the only area in which a statistical agency was influenced by the policies of its departmental "parent." CIOs of Federal departments are charged with carrying out many of the privacy and security requirements of the E-Government Act. In addition, technology advances made it possible to think of managing information security risks and achieving cost efficiencies by standardizing and consolidating previously disparate information technology systems. As a result, statistical agencies were the recipients of departmental guidance that of necessity had to apply to widely diverse agencies. Sometimes the guidance layered on reporting and other requirements within the department, building off of what the law or OMB guidance might require. IT consolidation in particular seemed to raise difficult questions for data protection, since the access control shifted from statistical agencies to departmental CIOs.

This discussion has elaborated on why statistical agencies' data protection efforts do, in some respects, position them to implement OMB breach notification policy. It also makes clear that statistical agencies are often far from the lone actor in carrying out a statistical collection. They often are managing a variety of contractors, subcontractors and employees who are often geographically dispersed and connected "virtually" via technology. In addition, they often are carrying out a variety of internal requirements for reporting and other activities to comply with departmental requirements. In this environment, how have statistical agencies responded?

4.3 Implementing Recent Additions to OMB Information Security Policy

Statistical agencies, like other agencies, have responded to the OMB requirements in a variety of ways, both within and across agencies. One or more of the principal statistical agencies have taken the following steps which will help it comply with the new OMB guidance:

- Inventory and review uses of PII, particularly of Social Security numbers (SSNs);

[9] For example, Zarate, Alvan O. "General Principles in Establishing Effective Confidentiality Practices" prepared for the Joint Statistical Meetings, August 11, 2002.
[10] http://www.fcsm.gov/committees/cdac/cdac.html

- Prepare and develop compelling justifications for continued use and retention of PII, especially SSNs;
- Examine safeguards for continued necessary use of PII and SSNs;
- Extending accountability for reporting breaches and other specific requirements to contractors and subcontractors by creating standardized contract language;
- Preparing practically via procedures and messaging to respond in the inevitable occurrence of a potential breach; and
- Determining how to evaluate risk from loss, especially of indirect identifiers, to help determine when to respond and how.[11]

Recognizing the importance of these issues, in the spring of 2007, the Federal Committee on Statistical Methodology (FCSM) formed a subcommittee to focus on privacy issues affecting statistical agencies. The privacy subcommittee has provided an opportunity for the statistical agencies to share their lessons learned and status of the activities listed above. The privacy subcommittee has also provided an opportunity to reach across agency lines to identify and consider solutions to issues that are affected by common organizational structure issues. In fact, one of the subcommittee's major emphases since its creation in 2007 has been working together to implement a standardized interpretation of reporting and notification practices. This standardization has been facilitated by interaction with OMB's Privacy Act lead, but is also limited to a degree by the customized requirements and guidance issued by individual departments.

These discussions have led to a better understanding of the different statutory responsibilities statistical agencies face as statistical agencies (i.e., CIPSEA or other confidentiality statutes) and as components of larger Federal departments or agencies covered by FISMA, the responsibilities for which are typically assigned to the CIO. The respective responsibilities can result in different starting points in understanding various terms and concepts. The different points of view can require significant collaboration, including, for example, meeting requirements from CIOs that appear at times to be not fully consistent with statistical agency confidentiality obligations.

For example, in the area of IT consolidation, statistical agencies have put great effort into communicating to CIOs their perspective that CIPSEA and other similar confidentiality pledges do not allow for data access to departmental individuals even within an IT office unless they are formal agents with appropriate training. The statistical agencies see data control as something most readily achieved by maintaining direct control of the computers and servers on which the statistical data reside. CIO's responsibilities to secure all systems and data across a department in an efficient manner tend to require flexibility in who manages which machines and in moving data around on them. In this perspective, all data are secure but control is occurring at the departmental CIO level, not at the statistical agency level.

Additionally, the "one hour" reporting rule for breach notification has sometimes been interpreted more stringently to require the statistical agency to report internally within 30 minutes, and at times has been understood to be from the time the breach was discovered even by a subcontractor. Such a perspective can lead either to inevitable delays because the contractor process relies on a "chain of command" reporting process or to on-time reporting that skips the chain of command and causes communications difficulties in the aftermath.

Such discussions among the FCSM group have led to a better understanding of the value of cultivating a "common language" and a desire to better understand the requirements and demands on CIOs, in order to identify more mutually satisfactory strategies for achieving the jointly held goal of protecting agency-held information.

5. Looking Ahead

Statistical agencies recognize that research and further discussions at the agency level will continue to be important, both in the area of breach notification and with respect to Federal information security and privacy, as each individual statistical agency determines the processes and practices it needs to meet its current responsibilities. One pressing research need is to better understand the effect of these policy requirements, particularly as they may affect participation in voluntary surveys. Does greater transparency reassure respondents, scare them, or have no effect? The

[11] The final item on the list is the focus of the Seastrom paper presented in the same session.

Seastrom paper is an early look at possible effects on survey participation, particularly the negative implications of breach notification where risk is minimal. However, the Seastrom paper reflects a breach incident prior to the final issuance of Memorandum 07-16, and describes experiences with only one survey subpopulation. Additional agency research could give greater insight into optimal strategies for implementing the policy in a manner most responsive to respondents' needs while maintaining the statistical viability of affected surveys.

Statistical agencies could also benefit from considering enhanced communication mechanisms to facilitate an ongoing dialogue with departmental CIOs. The goals could be several. First, statistical agencies could explore how each entity's requirements work in tandem to avoid rework or double the work. Statistical agencies could focus on ensuring that the final product is an effective set of policies, procedures and controls for protecting respondent data. Finally, the communications could identify forward-looking mechanisms to build in flexibilities to adapt to new requirements and risks without great cost and disruption.

Last but not least, consistent with the emphasis in Memorandum 07-16, statistical agencies could ensure that their attention to data breach responses does not give way to the risk of unduly shifting focus from long-standing requirements to prevent breaches in the first place. Working across statistical agencies and departments can help on all of these fronts, and is therefore an important ongoing focus.

Acknowledgements

The authors gratefully acknowledge the Federal Committee on Statistical Methodology's Subcommittee on Privacy for its willingness to allow us to participate in its discussions and to cite it for this paper. We also acknowledge our colleagues at the Office of Management and Budget for their helpful review and guidance, especially Katherine Wallman, Kimberly Nelson, Jasmeet Seehra, Paul Bugg, Kristy Daphnis and Brian Harris-Kojetin.

References

Confidentiality and Data Access Committee (CDAC) webpage, last accessed September 24, 2008 (http://www.fcsm.gov/committees/cdac/cdac.html).

"Personal Privacy in an Information Society: The Report of the Privacy Protection Study Commission," transmitted to President Jimmy Carter on July 12, 1977.

Pulliam, Daniel. "Education data breach puts 21,000 student loan borrowers at risk," in *Government Executive*, August 24, 2006 (http://www.govexec.com/dailyfed/0806/082406p1.htm).

U.S. Office of Management and Budget. "Annual Report to Congress on Statistical Programs of the United States Government, Fiscal Year 2008," September 14, 2007 (http://www.whitehouse.gov/omb/inforeg/statpolicy.html#sp).

U.S. Office of Management and Budget. "Implementation Guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)," in the Federal Register, Vol. 72, No. 115, page 33362, June 15, 2007 (http://www.whitehouse.gov/omb/fedreg/2007/061507_cipsea_guidance.pdf).

VanBokkelen, James B. "2006: The year of the breach--Network forensics will play a bigger role in the discovery and remediation of data breaches," in Federal Computer Week, December 18, 2006 (http://www.fcw.com/print/12_44/news/97098-1.html).

Zarate, Alvan O. "General Principles in Establishing Effective Confidentiality Practices" prepared for the Joint Statistical Meetings, August 11, 2002.