Comments of the World Privacy Forum

Regarding

Office of Management and Budget's Proposed Implementation Guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)

December 15, 2006

Brian Harris-Kojetin, Ph.D
Office of Information and Regulatory Affairs
Office of Management and Budget
725 17th Street, NW
New Executive Office Building, Room 10201,
Washington, DC 20503

VIA email to bharrisk@omb.eop.gov

This is a comment on OMB's Proposed Implementation Guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). The request for comments appeared in 71 Federal Register 60772 (October 16, 2006).

These comments are submitted by and on behalf of the World Privacy Forum, a non-profit, non-partisan public interest research organization. The World Privacy Forum focuses on in-depth research and analysis of privacy topics. See <http://www.worldprivacyforum.org>.

A. USA-Patriot Act

A major deficiency with the proposed guidance is its failure to address the relationship between CIPSEA and the USA PATRIOT Act. Section 215 of the USA PATRIOT Act (50 U.S.C. 1861) gives broad authority to the FBI to demand access to records. As an after-enacted statute, CIPSEA appears to limit that authority in the case of information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes. While the WPF believes that CIPSEA's restrictions prevent disclosures under the USA PATRIOT Act, other interpretations of the law are possible. The absence in the CIPSEA guidance of any discussion of the USA PATRIOT Act is troubling. The guidance should state expressly that exclusively statistical data is beyond the reach of the USA PATRIOT Act. The absence of a statement to this effect will raise concerns and suspicions that the protections of CIPSEA are less than what initially meets the eye.

Questions about the relationship between CIPSEA and other laws – preexisting or after-enacted – are certain to arise in the future. The guidance should include a definitive statement by the Office of Management and Budget that CIPSEA's non-disclosure policy trumps any other existing statute or regulation – including disclosures under Section 215 of the USA PATRIOT Act, disclosures in response to grand jury subpoenas, disclosures under other forms of compulsory process, and any other disclosures not expressly provided for in CIPSEA. A clear statement will reassure the public, diminish the possibility that an agency may someday reach a different conclusion, and save time and expense of duplicative policy and legal activities.

B. Identifiability

The discussion of the identifiability of personal information in the proposed guidance is insufficient. Identifiability is admittedly a difficult issue, and CIPSEA offers a definition that shows more sophistication than is usually found in other federal laws:

> (4) The term ''identifiable form'' means any representation of information that permits the identity of the respondent to whom the information applies to be reasonably inferred by either direct or indirect means.
>
> The proposed guidance merely says:
>
> For the purpose of this guidance, confidential information refers to any *identifiable* information, regardless of whether direct identifiers such as name and/or address have been removed from the individual records.

OMB can and should say more than this. We are aware of the reference in the guidance to Statistical Policy Working Paper #22 and to the Federal Committee on Statistical Methodology's Confidentiality and Data Access Committee's disclosure review checklist. These references are appropriate and helpful. We are also aware that identifiability is and has been a significant concern of statistical agencies for a long time. All of this is reassuring.

Nevertheless, both agencies and data subjects would benefit if there were more specific guidance about the meaning of the terms *reasonably inferred* and *direct or indirect means*. It would also be useful if OMB clarified how the CIPSEA standard specifically relates to the HIPAA standards of *no reasonable basis to believe* and *risk is very small* [45 CFR §164.514(a) & (b)(1)(i)] as well as to comparable standards found here and there in other federal laws. Agencies need guidance whether a risk assessment is required, how to conduct that risk assessment, what data sources (public and private) must be considered in assessing identifiability possibilities, how much effort and cost are reasonable, and answers to other questions pertaining to the standard for identifiability.

The ever-increasing pools of public and private data, advances in computing technology, and improvement in techniques for matching data make identifiability a moving target and an area of great concern. The realm of micro data – and even tabular data– that cannot be reidentified seems to shrink every day. OMB should provide more clarity and sharper lines.

C. Privacy Act of 1974

For records protected under CIPSEA that are also maintained in systems of records subject to the Privacy Act, the disclosure restrictions present several challenges not addressed in the proposed guidance.

First, subsection (b) of the Privacy Act of 1974 authorizes numerous disclosures, many of which are inappropriate for CIPSEA records. For example, disclosures for law enforcement purposes (except perhaps for violations of CIPSEA) allowed under (b)(7) appear to be improper. The disclosure for statistical purposes in (b)(5) appears to be wholly superseded by CIPSEA's requirements. Disclosures to Congress, pursuant to court order, and to consumer reporting agencies are inappropriate. Whether disclosure to GAO is consistent with CIPSEA may be a more debatable question. Even disclosures to officers and employees of the agency who have a need to know raise issues. CIPSEA appears to supersede the authority here as well by establishing clearer, more specific, and additional requirements.

Second, disclosures by routine use for systems containing CIPSEA records also need attention. Many of the familiar routine uses used by agencies are clearly inappropriate (e.g., law enforcement), need adjustment (e.g., agency contractors), or may require additional thought (e.g., NARA record management activities). Also, many agencies have established blanket, general, prefatory, or agency-wide routine uses that apply to all agency systems of records. The automatic application of these routine uses to CIPSEA systems is inappropriate and improper. Agencies that use agency-wide routine uses will need direction to deal with the conflicts. Continuing to apply routine uses to statistical systems of records when the routine uses are inconsistent with CIPSEA is a violation of the Privacy Act of 1974.

Agencies would benefit if OMB would elaborate on the intersection between CIPSEA and the Privacy Act of 1974. Agencies should be told how to address the subsection (b) conditions of disclosure, how to narrow or eliminate routine uses, and what to do with agency-wide routine uses. OMB has responsibility for providing guidance on the Privacy Act of 1974 as well as CIPSEA, so it should be able to address all of the issues with existing expertise.

D. Public Notice and Availability

Footnote 31 on page 11 tells agencies that they should consult with OMB regarding use of agents to perform exclusively statistical activities. The use of agents should be both fully transparent and the subject of public notice and comment. We suggest that in addition to consulting with OMB, agencies should be expressly told to provide public notice and to request public comment.

Part VIII of the proposed guidance (beginning on page 28) addresses annual reporting by agencies. The guidance should say that the reports must be public and posted on agency websites.

E. Confidentiality Pledges

Part II B of the proposed guidance addresses informing respondents. Notifying individuals is always a delicate issue that involves tradeoffs between competing goals and limitations. Ongoing debates about the length and content of privacy notices on websites and elsewhere illustrate the complexity of the subject.

The proposed guidance attempts to strike a balance between competing objectives. However, it is not clear that it succeeds as well as it might. For example, the guidance offers a shortened version of the confidentiality pledge using these words:

> The information you provide about *(choose one* -- yourself, household, establishment*)* will be used for statistical purposes only. In accordance with the Confidential Information Protection provisions in Public Law 107-347 (*option to add* and other applicable Federal laws*)*, your responses will be kept confidential and will not be disclosed in identifiable form *(optional* - without your consent).[42] By law, everyone working on this *(your agency here)* survey is subject to a jail term, a fine, or both if he or she discloses ANY information that could identify any confidential survey response.

We applaud the attempt to find a shorter version that might be used on the telephone or elsewhere. However, it is unlikely that any normal person would be able to ingest the proposed language on the telephone. The citation of a public law number will be incomprehensible to nearly everyone, and few would even have the ability to investigate and understand the law being cited. The short version is not particularly short either.

The World Privacy Forum suggests that OMB consider developing a formal statistical confidentiality seal controlled by a federal agency (possibly OMB). The purpose would be to provide an identifiable marker that would tell individuals if the information they provide will receive the highest degree of confidentiality protection available under law. A well-designed seal would also tell respondents when a lesser degree of protection is available.

The design of a seal would take some effort, and we do not pretend to have the expertise to offer a fully-conceived example here. One idea for consideration would use a green-yellow-red color scheme. Green would mean respond with confidence because answers receive the highest level of legal confidentiality protection. Yellow would mean respond with caution because answers receive some confidentiality protection but less than the highest level of legal protection. Red would mean no legal confidentiality protections at all.

The adoption of a recognizable and variable seal would, over time, help potential respondents learn about the consequences of providing information. It would also help to meet the goal stated on page 25 of the proposed guidance:

> Agencies must ensure that the public is able to distinguish easily between their CIPSEA pledge and any non-CIPSEA pledge covering information that will be used for non-statistical purposes.

A seal should also incorporate a web address where more information is available about the meaning of the protections. While a seal would not be a substitute for full disclosure of the elements identified in the proposed guidance, it would offer a shorthand and easily recognizable designation of the risk involved and would meet the needs of some potential respondents. Once developed, a statistical confidentiality seal would also help in developing language to be disclosed in telephone collection activities.

F. Miranda Warnings

Footnote 22 on page 8 of the proposed guidance discusses briefly the possibility that confidential responses may be used to report child abuse. This is an example of a disclosure of a confidential response that might result in the imposition of sanctions on the respondent. More attention needs to be focused on the possibility that an individual might provide a confidential answer yet encounter consequences notwithstanding the expectation of confidentiality created at the time of information collection.

For child abuse, it is appropriate to offer a reasonable warning in those circumstances in which it is foreseeable that the responses might expose abuse. In other circumstances, no warning may be needed. An institutional review board (or the equivalent) might help to decide if a warning is appropriate.

In other cases, a warning may be either insufficient or directed at the wrong person. The same example illustrates both cases.

Prior to the USA PATRIOT Act, education statistics collected by the National Center for Educational Statistics had strong statutory confidentiality protections against secondary use. The PATRIOT Act changed that law (later recodified at 20 U.S.C. § 9573) to make the information available for terrorism investigations.

For statistical information collected with the express statutory possibility of use during administrative or judicial proceedings, the equivalent of a limited Miranda warning is necessary. Respondents need to be told in an appropriate way that *anything you say can and will be used against you in a court of law*. This may be suitable for information that will result in child abuse reporting and with educational statistics subject to 20 U.S.C. § 9573. There may be other examples of laws that permit the use of statistical information for the non-statistical purpose of criminal prosecution or other sanction. Without a detailed and express warning, respondents may be placed in jeopardy without due process.

In the case of educational statistics, the information reported to the government may be obtained from a person (e.g., a school) other than the data subject. In this case, a warning to the respondent goes to the wrong person. There must be a way to provide a warning and a choice to the data subject. Otherwise, the warning is meaningless because it never reaches the individual who may be in peril. As long as identifiable data is being collected or if the collected data is subject to reidentification, warning and choice must flow down to the data subject through the institution that collects the information and conveys it to the government.

The solution to this problem is not likely to be simple and may vary from circumstance to circumstance. It is imperative, however, that it be addressed. If statistics are "laundered" through third parties in a manner that fails to provide a suitable warning and an effective choice to the data subject, then CIPSEA's goal of protecting individuals who provide information for statistical purposes cannot be achieved. When statistical information is fully protected by law from any secondary use against the data subject, concerns about warnings and decisions by data subjects are lessened and may disappear entirely. However, in those cases where a data subject retains exposure to criminal prosecution or other consequences, a timely and complete warning is necessary, along with an effective choice. We suggest that OMB consider this problem and issue additional guidance.

G. Laptops

The proposed guidance addresses physical and information systems security (page 22). Recent events have highlighted the particular vulnerability of laptop computers to loss and theft. The consequences for data subjects (as well as data controllers) have also become well known. The proposed guidance would be enhanced if a sentence or two were added about the need for additional attention to the security of potentially identifiable statistical information maintained in laptops. If PDAs or other types of portable information technology are routinely used in some circumstances, then guidance about the need for securing personally identifiable information held on those devices should also be included.

Thank you for the opportunity to comment on the CIPSEA guidance.

Respectfully submitted,

Pam Dixon
Executive Director, World Privacy Forum
www.worldprivacyforum.org