Statement of Ari Schwartz
Deputy Director
Center for Democracy & Technology
before the
Committee on Homeland Security and Governmental Affairs
on E-Government

December 11, 2007

Chairman Lieberman, Ranking Member Collins, and members of the Committee, thank you for holding this hearing on E-Government. I am Ari Schwartz, Deputy Director for the Center for Democracy & Technology (CDT).

CDT is a non-profit public interest organization founded in 1994 to promote democratic values and individual liberties for the digital age. CDT works for practical, real-world solutions that enhance free expression, privacy, universal access and democratic participation. We are guided by our vision of the Internet as a uniquely open, global, decentralized and user-controlled medium. We believe the Internet has unprecedented potential to strengthen democracy and encourage citizen participation by placing powerful information and communications technology in the hands of individuals and communities.

The Role of the E-Government Act of 2002

For five years, the E-Government Act has promoted improvements in the federal government's use of information technology, and has resulted in government information resources being more readily available to the public. The law showed great foresight in focusing on issues such as accessibility of information, privacy and security, which remain of central concern to the public.

The principal role of the E-Government Act has been to promote best practices among agencies in important areas and to solidify the federal government's technology management structure. Unquestionably, the E-Government Act has changed the way that the public interacts with the government for the better. For instance, a citizen can look up pending regulations, corporate filings, and search the federal agency websites through the USA.gov portal, regulations.gov and other appropriate sites.

We have also learned a great deal from agency implementation of the law about what areas can be improved. Five years of experience, technological progress, and changes in user expectations should guide revisions to the E-Government Act to facilitate availability of public resources to the public, and privacy protections for new technologies.

Making Government Information Searchable

The Pew Internet Project has found that commercial search engines are the most popular means to find government information.[1] This is true for several reasons. First, citizens don't necessarily know which agency holds the information they seek, but they often know how to search for it. Also, commercial search engines have simply become the most efficient and effective route to find information online. Government agencies must recognize that taxpayers will not find the information that is made available unless this information can be found on commercial search engines. Some agencies have public information resources that are not immediately accessible via search engines due to relatively minor technical problems that the agencies should quickly remedy.

Today, the Center for Democracy & Technology and OMB Watch are releasing a report demonstrating the types of government information that are not available through search engines and why. The full report is attached as an Appendix to this testimony, but I will offer a quick summary of the most important points.

In order to find online information, commercial search engines continually index the Internet via simple programs called crawlers. These crawlers face certain technological limitations that often prevent them from indexing information. Luckily, there are relatively simple ways to make information more accessible to search engines, and help government sites ensure that the most relevant information is provided to the public. Two easy ways to ensure that government information is indexed are to adopt the Sitemaps protocol, which guides search engines to content, and to limit the use of robots.txt files, which ask search engines not to crawl certain content. Unfortunately, CDT and OMB Watch found many important federal government agencies offering information and services that were not being indexed for search because they did not use these protocols well. Select examples of information that cannot be fully found by citizens using commercial search engines include:

- Federal Emergency Management Agency databases: including Flood Map Modernization project at FEMA, which maps out flood hazards.
- Other Department of Homeland Security databases: including topics like environmental radiation monitoring.
- FedBizOpps.gov database: listing approximately 200 government business opportunities within the field of telecommunications.
- Central Contractor Registration database: listing who does business and receives moneys from the federal government
- Federal Procurement Data Services database: includes data on all government contracts, including all telecommunications contracts.
- Smithsonian Institute resources: including many online content collections, including the Smithsonian Institution Research Information System.
- National Oceanic and Atmospheric Administration databases: including databases used to monitor environmental data and research.

---

[1] John B. Horrigan, "How Americans Get in Touch with Government" Pew Internet Life, May 24, 2004 — http://www.pewInternet.org/pdfs/PIP_E-Gov_Report_0504.pdf

- Bureau of Labor Statistics databases: including many of the statistics and collections of information hosted on the BLS site.

It is unclear to CDT and OMB Watch whether these agencies know that their information is not publicly searchable and have not taken the adequate steps to change their practices or whether the agencies simply do not know that this important information is not being crawled. In either case, our findings show that this is a systematic problem that should be addressed as soon as possible.

It should also be noted that even the government's own search engine is directly impacted by this problem. The USA.gov site utilizes Microsoft Live Search to run its search capability. Therefore, it is subject to exactly the same inability to search these important sites as other commercial search engines.

Fortunately, the E-Government Act recognized the importance of the availability and accessibility of information. Section 207 of the Act was meant to improve the organization and categorization of government information. OMB was directed to require that agencies proactively improve access to government information and services. As President Bush said in his signing statement for the E-Government Act, "[t]he Act will also assist in expanding the use of the Internet and computer resources in order to deliver Government services, [...] for a citizen-centered, results-oriented, and market-based Government."[2] Recently, by passing the searchability provision of the Reauthorization of the E-Government Act, this Committee helped to ensure that this provision was modernized to include promote best practices that could be used to tackle this problem.

We urge the Committee to work with us to encourage agencies that have not made public information available to search engines to do so immediately and to oversee proper implementation of the search provisions of the Reauthorization Act to ensure prompt compliance.

Privacy Impact Assessments

The increased ability to find information brings with it the challenge to better manage, protect and secure the personal information of individuals held by government that could inadvertently be made public if proper steps are not taken. Congress clearly understood this concern when it passed the E-Government Act. Section 208 of the Act was specifically designed to "ensure sufficient protections for the privacy of personal information as agencies implement citizen-centered electronic Government."[3] The method to achieve this goal was to increase transparency about how the government collects, manages and uses personal information about individuals through Web privacy notices and privacy impact assessments (PIAs).

---

[2] http://www.whitehouse.gov/omb/egov/g-3-statement.html
[3] PL 107-347, Section 208

The E-Government Act required that agencies perform PIAs before adopting new technology or using collections of personally identifiable information. These PIAs are public documents, containing a description of the project, a risk assessment and a discussion of potential threats to privacy, and ways to mitigate those risks. PIAs ensure that universal privacy concerns are considered as part of these decisions, and that the public has access to this element of the decision making process.

Over the past five years, PIAs have become an essential tool to help protect privacy. They are sometimes called "one of the three pillars" of the US government privacy policy.[4] Unfortunately, as with the other privacy laws, the federal government has unevenly implemented even the basic transparency requirement of PIAs across agencies.

The guidance issued by OMB pursuant to the Act with respect to PIAs was vague and has simply not provided agencies with the tools they need to successfully implement the PIA requirement unless they already had privacy experts on staff. While some agencies, like the Department of Homeland Security (DHS),[5] have set a high standard for PIAs and have continued to improve them over time, the lack of clear guidance has led some agencies to create cursory PIAs or none at all. For example, despite the major privacy implications of the use of RFID in passports, the US Department of State gave the issue only cursory consideration in its PIA, a document of only ten sentences.[6] Even more troubling is the finding that some agencies simply do not perform PIAs on as many as half their qualifying technologies.[7] Other agencies, even those that prepare in depth PIAs, too often complete them after a project has been developed and approved. PIAs are supposed to inform the decisionmaking process, not ratify it.

The inconsistent implementation of PIAs should be of great concern to this committee. The work of the agencies that have taken the mandate to develop PIAs seriously and used them as a tool for analysis and change should be used as a starting point for developing best practices for all federal agencies. CDT hopes that the provision included in the E-Government Act Reauthorization bill that passed out of this committee last month that

---

[4] DHS Chief Privacy Officer Hugo Teuffel, *Presentation before the European Commission's Conference on Public Security, Privacy and Technology*, November 20, 2007 Brussels, Belgium. Mr. Teuffel suggested that the three current pillars are the Privacy Act of 1974, Section 208 of the E-Government Act and the Freedom of Information Act.

[5] The DHS Website on Privacy Impact Assessment offers a range of resources to DHS components and to other agencies — http://www.dhs.gov/xinfoshare/publications/editorial_0511.shtm

[6] http://foia.state.gov/SPIAS/20061.DOS.PIA.Summary.Passport-cleared.pdf Also see CDT's letter May 2, 2007 letter to Secretary of State Rice on the agencies failure to provide adequate PIAs for this and a related project — http://www.cdt.org/security/identity/20070502rice.pdf

[7] OMB FY2006 Report to Congress on Implementation of the Federal Information Security Management Act of 2002, at www.whitehouse.gov/omb/inforegreports/2006_fisma_report.pdf

would specifically requires OMB to create best practices for PIAs across the government will help to address these problems.

As the Government Accountability Office and others have pointed out, OMB has not provided real leadership on privacy issues.[8] Best practices on PIAs can be a starting point for OMB to begin providing such leadership.

Even then, the transparency provided by PIAs must not be viewed as a full solution. Congress needs to begin to address more fundamental privacy issues within government agencies to ensure the trust of the American people. This should begin with a review of the Privacy Act of 1974 and a look into whether the law is adequate to address how the federal government today is using personal information. We look forward to working with this committee to help address these critical privacy issues in more detail in the near future.

Conclusion

The five years of experience in implementing the E-Government Act has provided valuable lessons in how to move government information services forward. In the short term, changes in the way people use the Internet mean that public government information online must be made accessible to search engines. Privacy impact assessments can be improved across the federal government based on the good work that has been done. In the long term, we will need leadership from OMB to protect privacy and security of Americans. We urge this committee to continue its leadership in adapting policy to fit the changing landscape and in oversight of that policy.

---

[8] Government Accountability Office, *Privacy Act: OMB Leadership Needed to Improve Agency Compliance*, June 2003 — http://www.gao.gov/new.items/d03304.pdf . Also see Kenneth A. Bamberger and Deirdre K. Mulligan, Privacy Decisionmaking in Administrative Agencies, University of Chicago Law Journal (forthcoming). A draft is available at— http://www.law.uchicago.edu/Lawecon/events/bamberger.pdf

Appendix: Center for Democracy and Technology and OMB Watch report
Hiding in Plain Sight: Why important Government Information Cannot Be Found Through Commercial Search Engines