



UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF INSPECTOR GENERAL

THE INSPECTOR GENERAL

MEMORANDUM

DATE: FEB 20 2001

TO: Craig Luigart, Chief Information Officer
U.S. Department of Education

FROM: Lorraine Lewis *Lorraine Lewis*

SUBJECT: Audit Report
Audit of the Collection of Personally Identifiable Information Through ED Internet Sites (Control Number ED-OIG/A11-B0002)

This Final Audit Report presents the results of our audit of the Collection of Personally Identifiable Information Through U.S. Department of Education (ED) Internet Sites. In your response to the draft report, you concurred with the basic findings in the report and summarized corrective actions completed and planned. Your full response is included as Attachment 3.

The Treasury and General Government Appropriations Act, 2001 (Act) signed into law on December 21, 2000 requires that not later than 60 days after the date of enactment, each agency's Inspector General shall submit to Congress a report that discloses any activity of the applicable department or agency relating to:

- (1) the collection or review of singular data, or the creation of aggregate lists that include personally identifiable information about individuals who access any Internet site of the agency; and
- (2) agreements with third parties to collect, review or obtain aggregate lists or singular data containing personally identifiable information relating to any individual's access or viewing habits for government and non-governmental Internet sites. (Title VI, Section 646).

AUDIT RESULTS

ED Activities Which Collect Personally Identifiable Information

We identified 54 ED activities which voluntarily collect personally identifiable information through the Internet. The information collected generally includes names, mailing addresses, email addresses, and phone numbers. Attachment 1 provides a list of the activities identified.

Use of Agreements with Third Parties

ED officials informed us that ED does not have agreements with third parties for the purpose of collecting personally identifiable information as stated in Title VI, Section 646 of the Act. During our limited testing of ED Internet sites and Web pages, nothing came to our attention to indicate that such agreements exist.

Findings on ED's Management of Internet Activities

During our audit, we reviewed 64 ED Internet servers, 111 Internet sites, and 4,056 Web pages that were on-line as of January 26, 2001. We analyzed ED's use of cookies on its Internet servers. A "cookie"¹ is information that a Web site puts on the user's computer so that it can track information about that user. There are two types of cookies:

- Session Cookie: A cookie used to retain and correlate information about users during a single session that expires when the user ends that browser session.
- Persistent Cookie: A cookie with a set expiration date that can be used to trace the activities of users over time. It may store the user's log-in information including password and email address used by the user to access an Internet site.

We designed our test to determine the existence of persistent cookies which would allow ED to collect and use personally identifiable information. We also tested whether ED adequately posted privacy policies on sites collecting personally identifiable information and tested whether ED used means other than cookies to collect personally identifiable information without the end user's direct knowledge.

As a result of these tests, we identified three areas needing additional oversight. Specifically, we found that ED needs to 1) strengthen controls over the use of persistent cookies, 2) ensure that privacy policy notices are provided, and 3) monitor methods for collecting personally identifiable information.

¹Definitions for "cookie" and other Internet technology terms are provided in Attachment 2.

Finding 1: ED Needs to Strengthen Controls Over the Use of Persistent Cookies

We found that 4 of ED's 64 Internet servers attached persistent cookies. In three occurrences, management was not aware that the servers were attaching persistent cookies. In the fourth occurrence, management was aware of the cookie but did not know that it was a persistent cookie lasting 36 years. While our limited testing did not identify that these cookies were collecting personally identifiable information, ED needs to ensure that controls are in place to detect the use of persistent cookies.

Office of Management and Budget (OMB) Memorandum 00-13 states "cookies should not be used at Federal web sites, or by contractors when operating web sites on behalf of agencies, unless, in addition to clear and conspicuous notice, the following conditions are met: a compelling need to gather the data on the site; appropriate and publicly disclosed privacy safeguards for handling of information derived from cookies; and personal approval by the head of the agency." The four servers we identified are:

- ocfo.ed.gov (also accessible as gcs.ed.gov) (Figure 1): This server attached a persistent cookie with an expiration date of December 30, 2010. The responsible ED official was not aware that this server was attaching persistent cookies and stated that any information that might be stored by these cookies was not being used by ED.

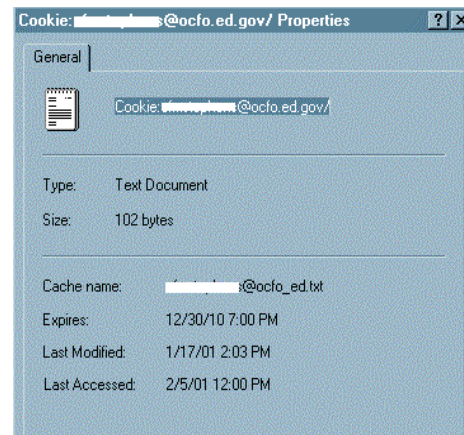


Figure 1. ocfo.ed.gov persistent cookie

- easitest.ed.gov (Figure 2): This server attached a persistent cookie with an expiration date of December 30, 2010. ED officials stated that this server was obsolete and should have been taken off-line about 18 months ago. These officials were not aware that the server was attaching persistent cookies and were unaware of any information that might be stored by the persistent cookies or how such information might be used.

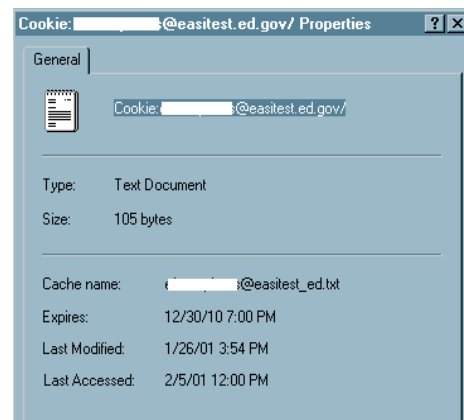


Figure 2. easitest.ed.gov persistent cookie

- *nle2.ed.gov* (Figure 3): This server attached a persistent cookie with an expiration date of September 26, 2037. ED officials were not aware that the server was attaching persistent cookies. The site has an average of over 18,800² unique visitors monthly.

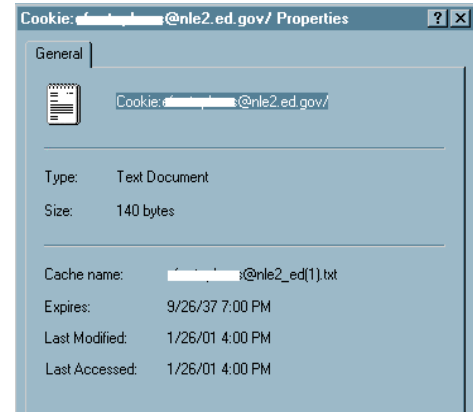


Figure 3. nle2.ed.gov persistent cookie

- *students.gov* (also accessible as *www.students.gov*)(Figure 4): This server attached a persistent cookie with an expiration date of September 26, 2037. ED officials were aware of this cookie but thought it was “temporary” lasting only 48 hours.

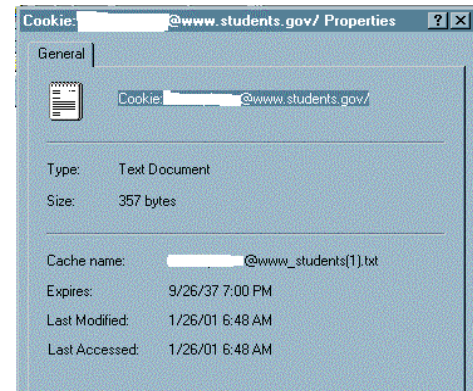


Figure 4. students.gov persistent cookie

ED officials told us that they were not aware of the presence of persistent cookies in the four servers we identified in our testing. Accordingly, none of these servers provided "clear and conspicuous notice" explaining the use of the persistent cookies to the end user as required by OMB 00-13. Three of the four servers identified did not have a privacy policy notice or a link to a privacy policy notice. The *students.gov* server had a link to a privacy policy, but it inappropriately described the cookie as a “temporary” rather than a persistent cookie.

In August 2000, the ED Internet Working Group³ provided the results of its survey on the use of cookies. ED Principal Offices were polled in July 2000 to identify all use of “cookies” on public access web sites in the *ed.gov* domain.⁴ ED officials identified 13 instances of session cookies. The survey identified one persistent cookie but this use met a valid business purpose and expired in 24 hours. The survey report stated that, "The survey indicates that ED is using cookies

² This number represents a three month average using the total number of unique visitors to the *nle2* server for the months of October, November, and December 2000.

³ The Internet Working Group is composed of principal office representatives who design processes for quality control and monitoring of ED's Web sites. It is chaired by the Office of the Chief Information Officer.

⁴ The *ed.gov* domain represents all ED controlled Web pages. We sampled about one percent of the Web pages for the domain. The persistent cookie identified by ED, ED Publications ordering, did not fall into our sample.

responsibly; no instances were discovered where cookie use was functionally unnecessary or accidental." Our results indicate that persistent cookies were being attached without management knowledge or a documented need.

ED's management controls for reviewing ED Web sites and World Wide Web operations did not specify that ED officials should review them for the use of persistent cookies. ED officials provided us with the internal policy, procedures, and checklists that they use for ED's Internet servers, sites, and Web pages. None of these documents provided instructions for reviewing existing servers regarding the use of persistent cookies.

Recommendations:

The Chief Information Officer should:

- 1.1. Determine whether the four persistent cookies identified in this report are necessary. ED should remove any unnecessary cookies and ensure that any remaining cookies comply with OMB 00-13.
- 1.2. Determine how the four persistent cookies identified in this report were attached without management knowledge and implement procedures to prevent future instances.
- 1.3. Revise existing procedures to require principal office officials to review servers, sites, and Web pages in advance to determine that unnecessary persistent cookies are not used.
- 1.4. Develop and implement procedures to periodically review all servers, sites, and Web pages to identify existing cookies and to determine that applicable legal and regulatory requirements have been met for their use.
- 1.5. Develop and implement procedures to remove obsolete servers, sites, and Web pages.

Finding 2: ED Needs to Ensure Privacy Policy Notices Are Provided

ED officials did not consistently provide required privacy policy notices for Web pages that collect personally identifiable information. We found that 32 (59%) of the 54 activities identified as collecting personally identifiable information through the Internet did not have privacy policy notices or links to notices. Privacy policy notices or links to these notices may have been used on "Home" Web pages, but they were not always used on the pages that collect personally identifiable information.

OMB Memorandum 99-18, *Privacy Policies on Federal Web Sites*, specifies that Federal agencies should provide privacy policy notices at major entry points, as well as at any web page where the agencies collect substantial personal information from the public. The memorandum adds that "each policy must clearly and concisely inform visitors to the site what information the

agency collects about individuals, why the agency collects it, and how the agency will use it. Privacy policies must be clearly labeled and easily accessed when someone visits a web site." Additionally, ED's *World Wide Web Server Policy and Procedures* (June 1999) specify that ED officials must review Internet sites in advance to ensure that information contained and transmitted via the ED-WWW Server will be secure, that all relevant Federal statutes are adhered to, and appropriate warnings, disclosures and/or disclaimers are openly displayed.

ED officials originally identified 46 activities where personally identifiable information is collected through the Internet. The activities collect personally identifiable information such as names, addresses, email addresses, and phone numbers. During our audit, we identified an additional 8 activities. Of these 54 activities, privacy policy notices were not provided for 32 (59%) of the activities as required by OMB and ED policies. Attachment 1 provides the 54 activities and identifies those that did not provide privacy policy notices.

For example, as illustrated in Figure 5, the Office of Elementary and Secondary Education's Reading Excellence Program has a registration form available on-line. At the time of our review, this Web page did not provide a privacy policy notice or a link to a notice. OMB specifies that privacy policy notices must be clearly labeled and easily accessed.

The screenshot shows a web browser window with a blue sidebar on the left and a white main content area. The sidebar contains the following navigation links: "News & Events", "Resources", "Sitemap", and "Home" (with an "OESE" logo). The main content area features a registration form titled "Reading Excellence and Class Size Reduction Grants Technical Assistance and Application Workshops Registration Form". The form includes the following fields:

- Name:
- Title:
- Name of Organization:
- Address:
- City:
- State:
- Zip Code:
- Telephone:
- Fax Number:
- E-mail:

Figure 5: Illustration of personally identifiable information collected through ED Internet site <http://www.ed.gov/offices/OESE/REA/application.html>

Recommendations:

The Chief Information Officer should:

- 2.1. Identify all major entry points to ED's Internet information as well as any web pages where ED collects personal information from the public. Once all of these locations are identified, ED needs to create privacy policy notices or obvious links to privacy policy notices as required by OMB.
- 2.2. Develop and implement management controls designed to ensure that privacy policy notices are located on needed entry points and Web pages of both existing and future sites.

Finding 3: ED Should Monitor Methods for Collecting Personally Identifiable Information.

After completion of our testing for cookies, we performed several analyses of the source code of the Web pages in the sample. This testing was performed to provide assurance that ED was not making use of Web Bugs, hidden forms, client side scripts, links to anonymous File Transfer Protocol (FTP) servers, or other means to collect personally identifiable information from Internet sites without a user's direct participation and knowledge. The results of our testing disclosed nine pages where ED used links to anonymous FTP servers.

- Five pages were on a site used by ED employees to read their government email over the public Internet.
- Four pages were on a site used to obtain technical reference documents related to Student Financial Assistance (SFA).

Web browsers have a built-in function that passes the user's email address to anonymous FTP servers when files are downloaded. Typical Internet users are often not aware that this type of personally identifiable information is being provided to the server. Our discussions with ED personnel responsible for managing ED's Internet sites indicated that they did not know that they were still running anonymous FTP servers.

Recommendations:

The Chief Information Officer should:

- 3.1. Determine if there is still adequate justification for maintaining anonymous FTP servers. If there is such a need, provide adequate disclosure to users that their email address may be collected using the server and how ED will use the email addresses collected.
- 3.2. Develop and implement procedures to identify and prevent Web Bugs, hidden forms, embedded client side scripts, links to anonymous FTP servers, or other means to collect personally identifiable information from Internet sites without a user's direct knowledge.

Auditee's Response and Auditor Comments

We received and reviewed your comments dated February 15, 2001. In your response, you state that the Office of the Chief Information Officer (OCIO) concurs with the basic findings of our draft audit report and that the report was complete and well done. Your response, which is included as Attachment 3 to this report, also contained observations and corrective actions completed and planned. In general, we find the draft report responses satisfactory to begin addressing our recommendations. The following paragraphs are additional comments that we have to specific statements in your response.

Your response states that the four persistent cookies we identified have been removed and that *easitest.ed.gov* has been removed from service. As of February 16, we found that the *easitest.ed.gov* server was still accessible and that the *students.gov* server was still attaching a persistent cookie. Other than testing for the removal of persistent cookies on these four servers, we have not taken steps to confirm that your stated completed actions were implemented.

Your response also states that SFA officials "disagree with the interpretation of what constitutes a site of interest" in reference to the additional sites or pages that we found as collecting personally identifiable information in Attachment 1 of this report. We used *OMB Memorandum 99-18 Attachment* as our criteria for information on what constitutes personally identifiable information. This memorandum includes email addresses as personally identifiable information. Although SFA officials disagreed with this interpretation, they agreed to modify the sites we identified to ensure that privacy policy notices are provided.

Additionally, your response included a clarification provided by the National Center for Education Statistics (NCES) officials concerning NCES Web pages identified in Attachment 1. We have updated Attachment 1 to include the NCES pages specified and noted that these pages have privacy policy notices.

BACKGROUND

ED uses a decentralized approach for ensuring that ED Internet sites and Web pages comply with laws, memoranda, and policies and procedures. Each Principal Office within ED is responsible for its respective Internet sites and Web pages and the content on the sites and pages. ED's Internet Working Group has a representative from each ED Principal Office. It is chaired by the OCIO. These members are involved in development efforts for their respective Principal Offices. All of ED, including SFA, are required to follow *World Wide Web Server Policy and Procedures* (June 1999) issued by OCIO.

AUDIT OBJECTIVES, SCOPE AND METHODOLOGY

Audit Objectives

The objectives of our audit were to report to Congress any activity of ED relating to:

- (1) the collection or review of singular data, or the creation of aggregate lists that include personally identifiable information about individuals who access any Internet site of ED; and
- (2) agreements with third parties to collect, review, or obtain aggregate lists or singular data containing personally identifiable information relating to any individual's access or viewing habits for government and non-governmental Internet sites.

Scope and Methodology

To fulfill these objectives, our audit focused on disclosing any activities that collect personally identifiable information through the Internet and disclosing agreements with third parties to collect personally identifiable information relating to any individual's access or viewing habits on ED or non-governmental Internet sites. For the activities that collect personally identifiable information, we tested to determine whether or not these activities had required privacy policy notices. We also tested our sample for the existence of persistent cookies and the use of Web bugs, hidden forms, embedded client-side scripts, links to anonymous FTP servers, or other means for Internet sites to collect personally identifiable information without a user's direct participation.

We interviewed officials from ED's OCIO and from the three offices responsible for the Internet sites or Web pages that assigned persistent cookies, specifically officials from SFA, Office of the Chief Financial Officer, and NCES.

Our audit included ED Internet servers, sites, and Web pages that were on-line as of January 26, 2001. Because ED continuously updates its Internet servers, sites, and Web pages, some of the information that we included in our review may not be currently available.

We used Internet search engines to identify web pages for each of the 111 Internet sites in our sample. We selected a sample of Web pages to review. When the site had fewer than 100 pages, we reviewed all of the pages. For sites having 100-199 pages, we reviewed the first 100. For sites having greater than 200 pages, we randomly sampled 100 pages.

We analyzed:

- 4,056 Web pages or approximately 1 percent of the total number of ED's Web pages;
- 64 ED Internet servers (100 percent of ED Internet servers available);
- 111 ED Internet sites (100 percent of ED Internet sites available); and
- 54 activities that collect personally identifiable information through ED Internet sites.

To perform our analysis, we used a software package to test ED's Internet servers for the use of persistent cookies. All other analysis was done by our auditors accessing ED Internet servers, sites, and Web pages; by reviewing ED documents; and by interviewing ED officials.

We performed our audit work at ED between December 8, 2000, and February 6, 2001. Our audit was performed in accordance with Government Auditing Standards appropriate to the scope of the review.

STATEMENT ON MANAGEMENT CONTROLS

As part of our audit, we assessed the system of management controls, policies, procedures, and practices applicable to our objectives. These included controls over the operation of ED's Internet sites. Because of inherent limitations, a study and evaluation made for the limited purposes of disclosing any ED activities that collect personally identifiable information through the Internet and disclosing ED agreements with third parties to collect personally identifiable information relating to any individual's access or viewing habits on ED or non-governmental Internet sites would not necessarily disclose all material weaknesses in management controls. However, our assessment identified management control weaknesses as discussed in the audit results section of this report.

ADMINISTRATIVE MATTERS

Please provide us with your final response to each open recommendation within 60 days of the date of this report indicating what corrective actions you have taken or plan to take and the related milestones.

In accordance with OMB Circular A-50, we will keep this audit report on the Office of Inspector General (OIG) list of unresolved audits until all open issues have been resolved. Any reports unresolved after 180 days from the date of issuance will be shown as overdue in the OIG's Semiannual Report to Congress.

Accordingly, please provide the Supervisor, Post Audit Group, Financial Improvement and Post Audit Operations, OCIO and OIG's Assistant Inspector General for Audit Services with semiannual status reports. These reports should address promised corrective actions until all such actions have been completed or continued follow-up is unnecessary.

In accordance with the Freedom of Information Act (Public Law 90-23), reports issued by OIG are available, if requested, to members of the press and general public to the extent information contained therein is not subject to exemptions in the Act.

We appreciate the cooperation given during the audit. If you have any questions or wish to discuss the contents of this report, please call Jack Rouch, Director, Systems Internal Audit Team on 202-260-3878. Please refer to the control number in all correspondence relating to this report.

Attachments

ED ACTIVITIES

#	ED IDENTIFIED ACTIVITY	PERSONALLY IDENTIFIABLE INFORMATION COLLECTED?		PRIVACY POLICY PROVIDED?	
		NO	YES	NO	YES
	www.ed.gov				
1	/comments/national/forum/question.html		1	1	
2	/comments/nationalforum97/index.html		1		1
3	/comments/problemform/ProblemForm.html		1		1
4	/comments/TeachLeader/index.html		1		1
5	/Family/agbts/tellus.html		1	1	
6	/Family/agbts_old/agbts98/form.html		1	1	
7	/Family/agbts_old/tellus.html		1	1	
8	/G2K/community/subscribe.html		1	1	
9	/inits/americareads/fwsform.html		1	1	
10	/inits/americareads/pcform.html		1	1	
11	/inits/Millennium/jazzreg.html		1	1	
12	/offices/OESE/REA/application.html		1	1	
13	/offices/OESE/t1.html		1	1	
14	/offices/OIG/feedback.htm		1	1	
15	/offices/OIG/hotlineform.htm		1	1	
16	/offices/OLCA/olcaform.html		1	1	
17	/offices/OPE/News/collegeweek/ ncwparticipate.html		1		1
18	/offices/OPE/thinkcollege/comment.html		1	1	
19	/offices/OPE/thinkcollege/early/about_us/ comment.html		1	1	
20	/offices/OUS/chip/pledge.html		1	1	
21	/offices/OUS/PES/NAVE/feedback.html		1		1
22	/Programs/EROD/EROD_collect.html		1	1	
23	/pubs/CompactforReading/survey.html		1	1	
	NCES				
24	National Postsecondary Aid Study NPSAS WEB (http://npsas.rti.org/)		1	1	
25	http://nces.ed.gov/newsflash/		1		1
26	National Assessment of Educational Progress (NAEP) Questionnaire		1		1
27	NAEP's My School http://nces.ed.gov/naep/myschool		1		1
28	NAEP Network List of state testing directors & NAEP coordinators.		1		1
29	surveys.nces.ed.gov/ipeds/		1		1
30	surveys.nces.ed.gov/library/als/		1		1
31	surveys.nces.ed.gov/library/stla/		1		1
32	Partnership for Family Involvement in Education Registration Form http://pfie.ed.gov "How to Join" link		1	1	
33	STW.ed.gov Website Comments & Suggestions stw.ed.gov/Database/comments.htm		1	1	

ED ACTIVITIES

34	Examples that work suggestion box www.stw.ed.gov/EXP_WORK.CFM		1		1	
35	STW Listserv www.stw.ed.gov/list2.htm		1		1	
36	e-application e-grants.ed.gov/e-App/eaHome.asp		1		1	
37	http://registerevent.ed.gov/ Teleconferences Registration Web Site		1		1	
38	ifap.ed.gov "Members Services"		1		1	
39	edwebenroll.ed. or sfawebenroll.ed.gov		1			1
40	fafsa.ed.gov		1			1
41	www.pellgrantsonline.ed.gov		1			1
42	www.loanconsolidation.ed.gov		1			1
43	www.dlserver.ed.gov		1		1	
44	edworkshop.walcoff.com/secure/main.htm		1		1	
45	www.ed.gov/DirectLoan/consolid2.html		1		1	
46	pin.ed.gov or eac.ed.gov		1			1
TOTALS		0	46		29	17

In addition to the sites identified by ED, we found 8 additional sites that collect personally identifiable information.

47	www.nsls.ed.gov		1		1	
48	www.afterschool.ed.gov		1			1
49	webx.ed.gov		1			1
50	test.ifap.ed.gov		1			1
51	sfablueprint.ed.gov / www.sfablueprint.ed. gov/sfa-vfa.ed.gov		1		1	
52	ombudsman.ed.gov / www.ombudsman.ed. gov/sfahelp.ed.gov		1			1
53	students.gov/www.students.gov		1			1
54	nle2.ed.gov		1		1	
TOTALS		0	8		3	5
GRAND TOTALS		0	54		32	22

Attachment 2**GLOSSARY**

1. **Browser:** An application program that provides a way to look at and interact with information on the World Wide Web.
2. **Client:** The requesting program or user in a client/server relationship. For example, the user of a Web browser is effectively making client requests for pages from servers all over the Web. The browser itself is a client in its relationship with the computer that is getting and returning the requested HTML file. The computer handling the request and sending back the HTML file is a server.
3. **Cookie:** Information that a Web site puts on the user's computer hard disk so that it can remember something about the client at a later time. (More technically, it is information for future use that is stored by the server on the client side of a client/server communication.) Typically, a cookie records the user's preferences when using a particular site. Using the Web's Hypertext Transfer Protocol (HTTP), each request for a Web page is independent of all other requests. For this reason, the Web server has no memory of what pages it has sent to a user previously or anything about the previous visits. A cookie is a mechanism that allows the server to store its own information about a user on the user's own computer.
 - **Persistent Cookie:** A cookie with a set expiration date that can be used to track the activities of the user over time. It may store the user's log-in information including password and email address used by the user to access an Internet site.
 - **Session Cookie:** A cookie used to retain and correlate information about a user during a single session. It expires when the user ends the browser session.
4. **FTP:** File Transfer Protocol, a standard Internet protocol that is the simplest way to exchange files between computers on the Internet.
 - **Anonymous FTP:** A method for giving users access to files so that they do not need to identify themselves to the server.
5. **Hypertext:** The organization of information units into connected associations that a user can choose to make. An instance of such an association is called a link or hyperlink.
6. **Internet Server:** In general, a computer program that provides services to other computer programs in the same or other computers. A computer that holds the files for one or more sites.
7. **Internet site:** A collection of Web pages on a particular subject that can be accessed.

8. **Link:** Using hypertext, a link is a selectable connection from one word, picture, or information object to another.
9. **Personally Identifiable Information:** Name, email address, Social Security number, or other unique identifier.
10. **Web bug:** A file object, usually an image file that is placed on a Web page or in an email message to monitor user behavior, functioning as a kind of spyware. Unlike a cookie, which can be accepted or declined by a browser user, a Web bug arrives as just another image on the Web page. A Web bug is typically invisible to the user because it is transparent (matches the color of the page background) and takes up only a tiny amount of space. It can usually only be detected if the user looks at the source version of the page to find an image that loads from a different Web server than the rest of the page.
11. **Web page:** On the World Wide Web, a page is a file notated with the Hypertext Markup Language (HTML). Usually, it contains text and specifications about where image or other multimedia files are to be placed when the page is displayed.



UNITED STATES DEPARTMENT OF EDUCATION

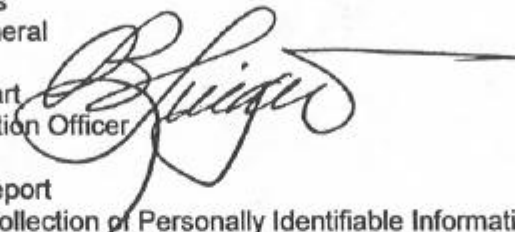
OFFICE OF THE CHIEF INFORMATION OFFICER

THE CHIEF INFORMATION OFFICER

FEB 15 2001

MEMORANDUM

TO: Lorraine Lewis
Inspector General

FROM: Craig B. Luigart
Chief Information Officer 

SUBJECT: Draft Audit Report
Audit of the Collection of Personally Identifiable Information Through ED
Internet Sites (Control Number ED-OIG/A11-B0002)

OCIO concurs with the basic findings of the subject draft audit report. Our comments which follow reflect not only the OCIO response but responses also from other parts of the agency which maintain internet sites – SFA, NCES, OCFO. We find the draft report complete and well done. We offer the following observations and report completed and planned corrective actions.

Finding 1: ED Needs to Strengthen Controls Over the Use of Persistent Cookies.

- 1.1 Determine whether the four persistent cookies identified in this report are necessary. ED should remove any unnecessary cookies and ensure that any remaining cookies comply with OMB 00-13.

See explanations under 1.2 below. The four persistent cookies were not necessary.

- 1.2 Determine how the four persistent cookies identified in this report were attached without management knowledge and implement procedures to prevent future instances.

ocfo.ed.gov The cookie was inadvertently established by the contractor when the site usage statistics software was installed. We didn't check for it, since there was no information being collected on the site other than the usual statistics about the number of users and most requested pages. The Web Trends software used to collect the site usage statistics has been adjusted to stop placing cookies.

easitest.ed.gov This server was to have been off-line a year and a half ago. It has now been removed from service.

nle2.ed.gov The OCIO contract monitor instructed the Internet Application Support contractor not to use cookies but neither party adequately verified. The

contractor inadvertently implemented persistent cookies by failing to modify default ColdFusion settings. Cookies have been removed from all but one of the approximately 15 ColdFusion applications. The remaining application will be converted within the next two weeks to use server-side session variables instead of cookies.

students.gov This was an anomaly. We believe that it was built on with old ColdFusion software. The cookie has been removed.

OCIO procedures will be modified to include cookie **tests as part of acceptance testing** checklist for future new applications and for modifications to ensure compliance. See responses 1.3 below.

- 1.3 Revise existing procedures to require principal office officials to review servers, sites and Web pages in advance to determine that unnecessary persistent cookies are not used.

OCIO will revise its web policies and procedures to require that all new applications and web page adds/changes must be reviewed for use of cookies and collection of individually identifiable information. Principal Offices other than SFA and NCES are not well-staffed or well-prepared to do this kind of review. Most of them don't do the basics like making sure their links work or their pages are accessible. For www.ed.gov, to ensure compliance, we'll need to do that in the Web Services Group (WSG) and contractor review of materials before they're posted to the live site. OCIO will require that the Web Services Group staff and Internet Application Support contract staff -the last two steps in the standard web posting workflow for ed.gov - consistently conduct such review on all postings. Principal Offices' editors will be made aware of the new policies and procedures via the Internet Working Group (IWG). OCIO procedures will require cookie tests as part of the acceptance testing checklist for future new applications and modifications to ensure compliance.

- 1.4 Develop and implement procedures to periodically review all servers, sites, and Web pages to identify existing cookies and to determine that applicable legal and regulatory requirements have been met for their use.

OCIO's initial research within this short comment period has not uncovered any automated tools for this purpose. Without automated tools, it would be excessively burdensome to manually check each of the more than 50,000 web pages on ed.gov on a periodic basis. OCIO proposes instead to:

Scan the site thoroughly once to ensure a clean baseline, using manual methods and concentrating on those items most likely to harbor cookies and collection of individually identifiable information, i.e., database applications, web discussion forums, CGI scripts, and embedded Javascript. Closely inspect future applications and web pages adds/changes to ensure continuing compliance.

- 1.5 Develop and implement procedures to remove obsolete servers, sites, and Web pages.

The ED Web redesign that is currently underway will institute a metadata schema that will include Archive/Update Action, Archive/Update Date, and Records Management elements to help identify web pages that need to be archived or removed. However, until the National Archives and Records Administration (NARA) issues clear guidance for web page retention schedules, there will continue to be uncertainty among Principal Office editors about whether to retain or remove aging pages.

Finding 2. ED Needs to Ensure Privacy Policy Notices Are Provided.

- 2.1 Identify all major entry points to ED's Internet information as well as any web pages where ED collects personal information from the public. Once all of these locations are identified, ED needs to create privacy policy notices or obvious links to privacy policy notices as required by OMB.

Twelve of the top twenty entry points, and eight of the top ten entry points, (based on January 2001 WebTrends reports) now link to ED's general privacy policy notice. Links will be added to the remaining eight entry points within the two next weeks.

In the absence of automated tools, OCIO will use manual methods and available aids such as the Inktomi site search engine and LinkScan link checking/site analysis software to help identify pages that are likely to collect personal information, e.g., all database applications, web discussion forums, CGI scripts, pages containing the words "last name," "first name," "email address," or "phone number". OCIO will conduct a one-time scan and cleanup to establish a clean baseline.

- 2.2 Develop and implement management controls designed to ensure that privacy policy notices are located on needed entry points and Web pages of both existing and future sites.

See response to 1.4.

Finding 3. ED Should Monitor Methods for Collecting Personally Identifiable Information.

- 3.1 Determine if there is still adequate justification for maintaining anonymous FTP servers. If there is such a need, adequate disclosure should be provided to users that their email address may be collected using the server and how ED will use the email addresses collected.

The anonymous FTP service on emai102.ed.gov has been shut down by disabling anonymous access on emai102.ed.gov and by blocking FTP access to emai102 through the ED firewall.

Web Bugs, that is hidden forms, controls, links to anonymous FTP sites: Our Standard Operating Procedures includes a review of content and imbedded controls before any pages are posted to the ed.gov sites. In addition, ed.gov is managed by the TeamSite application which continually compares production

content to backup copies and overwrites any unauthorized changes to production content. This has the effect of removing any malicious Web Bugs.

- 3.2 Develop and implement procedures to identify and prevent Web Bugs, hidden forms, embedded client side scripts, links to anonymous FTP servers, or other means to collect personally identifiable information from Internet sites without a user's direct participation and knowledge.
- To test our sites, the IG used a browser that they were able to control from a script. They brought up each page on their list, looking for cookies before and after the page was visited, and noted any cookies that were left on the client after visiting a page. This is a tedious resource consuming process.
- A primary concern is the resources that would be required to review the amount of content that is updated daily. We will evaluate alternative solutions but have reservations about using the sitemap-driven automated testing procedure since ed.gov contains more than 50,000 individual pages which would probably put a serious strain on both the server and client if a script ran each page through a browser. The procedures would require manual attention to each step in the process unless the page that produced each cookie were automatically logged. We are looking at the upcoming release of the LinkScan link-checking tool we already use to check the site each night to see whether it can be used in a more efficient way to do the same job.
- See also the responses to 1.4 and 2.1.

Attachment I. ED Activities. The following comments apply to the sites listed in Attachment I.

From SFA: Of the eight sites identified as "found" by the IG, five are SFA sites. The wording in the report suggests that SFA failed to identify these sites as collecting personal data. In fact, we disagree with the interpretation of what constitutes a site of interest. These sites have in common an HTML "e-mail" page for reporting problems, or requesting more information, into which the user types his/her email address. Since we do not maintain this information in any sort of information store, the site managers apparently believed they were not "collecting personal information." However, it is SFA policy to provide reminders and assurances of privacy protection wherever there is even the possibility of a question, so we will modify these sites accordingly.

From NCES: Item 29 of attachment 1 of the draft audit report is incorrect. The link indicated, ed.gov/surveys/, is an asp page that contains an overview of the various surveys conducted by NCES. It does not collect any personally identifiable information. Our surveys are collected online at the following URLs:

IPEDS:	< http://surveys.nces.ed.gov/ipeds/ >
Academic Libraries:	< http://surveys.nces.ed.gov/library/als/ >
State Library Agencies:	< http://surveys.nces.ed.gov/library/stla/ >

These sites do collect personally identifiable information. Each of these URLs is an introductory page describing the survey - complete with the OMB clearance number. Each page contains a privacy policy synopsis describing the site's use of cookies as well as a link to the full NCES privacy policy.

From OCIO: Twenty-four of the items listed as lacking the required privacy policy links (all items identified on www.ed.gov, pfie.ed.gov, and stw.ed.gov) have been fixed:

D 16 items were corrected by adding the required link to the general ED privacy policy statement and/or adding language specific to the page: #s 1, 5, 8, 9, 10, 14, 15, 16, 18, 19, 22, 23, 30, 31, 32, 33

D 6 items were obsolete and were deleted: #s 6, 7, 11, 12, 13, 20