IRS.GOV General Support System (GSS) (Modernization and Information Technology Services (MITS) – 28) – Privacy Impact Assessment

PIA Approval Date – Oct. 31, 2011

System Overview:

This privacy impact assessment (PIA) updates the Modernization and Information Technology Services (MITS) –28 IRS.GOV to note the change of the use of persistent cookies on the website. The implementation of the persistent cookies is based upon the White House Open Government initiative, OMB Memorandum 10–22, dated June 25, 2010, and titled "Guidance for Online Use of Web Measurement and Customization Technologies". MITS–28 serves as the underlying architecture hosting applications that support the www.irs.gov website. This website was created to be the online presence of the Internal Revenue Service (IRS). Neither tax return data nor Sensitive But Unclassified (SBU) data is accessible or otherwise used on this public portal.

The core functionality of the components and other key information on the IRS.GOV is as follows: website core functionality (store, display and allow the download of IRS publications), disseminate tax news and information, search, application servers, database servers, report servers and backup servers. The IRS.GOV website provides links to online applications residing on the Registered User Portal (RUP). By utilizing proper identification, the RUP applications allow inquiries regarding claimed credits, refund or return status, or other information regarding taxpayer accounts. It should be noted that the RUP is not part of the IRS Public Portal, and therefore is not covered by this PIA.

Systems of Records Notice (SORN):

- IRS 34.037--Audit Trail and Security Records System
- IRS 00.001--IRS Correspondence Files (including Stakeholder Relationship Files) and Correspondence Control

Data in the System

1. Describe the information (data elements and fields) available in the system in the following categories:

A. Taxpayer – Each visitor to the IRS.GOV website (i.e., web user / Taxpayer) will have the following information recorded:
   - Internet Protocol (IP) address is retrieved from that user's computer and stored for security purposes only (e.g., in support of law enforcement activities or investigations)
   - Time of request to access IRS.GOV
   - Request details (specific web page)
   - Transmission Control Protocol/Internet Protocol (TCP/IP) Packet details (in the case of the Intrusion Detection System (IDS) for tracking malicious packet attacks)
   - Operating system of user
   - Browser version
   - Domain name
   - Referring website
   - Persistent and Session cookie identifier (ID) (used for compiling page views and visits into a unique, but anonymous, user session)
   - Visitor Display Color Depth (e.g., 32 bits, 24 bits, etc.)
   - Visitor Screen Resolution (e.g., 1024 x 768 pixels, 800 x 600 pixels, etc.)
   - Number of Bytes received

Collection and analysis of this information in the aggregate will enable the IRS to enhance site performance, make design improvements, improve informational materials available on our website, and improve customer service overall. In addition, some of this information is crucial for security investigations and customarily will not be used to uniquely identify users across sessions of web usages.

B. Employee – users who perform routine maintenance on the system or users who access content management to create, manage, and delete content for the website will have their user ID logged.

C. Audit Trail Information – IRS.GOV GSS auditing monitors internal user workstation and log on/off activities. It also monitors system administrator and security administrator activities while in their specific roles. The audit logs have critical event information:
   - Type of event
   - Source of event
   - Time and date of event
   - User accountable for event (user ID, terminal ID)

   Access to audit logs is also restricted to only the appropriate individuals to prevent unauthorized deletion or change of audit events. However, privileged users are authorized to select relevant events to be audited. Audit reports can also be generated using specific criteria (e.g. user ID, terminal ID). (“Appropriate individuals” includes contractor employees; “privileged users” include contractors.)

   Operational system logging (e.g., user ID, action, changes made, date, and time) is enabled on individual router and switch devices; however, there is currently not a login server in place where audit trails are captured. These events are periodically reviewed by telecom personnel and are immediately reviewed when an event occurs. Contractors are considered “telecom personnel”

D. Other – IRS web pages hosted on the IRS.GOV website are a main source of data content. On pages where website visitors voluntarily request information, publications, refund status, or other information, an appropriate application–specific privacy statement is posted. Each statement informs the visitor of the information being requested; why it is being requested; how it will be used and maintained; and, the impact if the information requested is not provided. Each page of IRS.GOV provides a link to the IRS Web Privacy Policy as well as links to taxpayers’ rights under the Privacy Act and other privacy protection statutes. Departure Notices are available for all viewers when leaving an IRS site.

2. Describe/identify which data elements are obtained from files, databases, individuals, or any other sources.

A. IRS – Any change requests to update web pages and web content on the IRS.GOV website originate within the IRS.

B. Taxpayer – The information stored will be that of standard access logs on Web, Application, Firewall, Intrusion Detection System (IDS), and Database systems. This information is crucial for security investigations and customarily will not be used to uniquely identify users (i.e., web users / Taxpayers) across sessions of web usages. A combination of client–side JavaScript on SmartSource Data Collectors (SDCs) and persistent and/or session cookies will be used on all IRS.gov web pages, as well as by some of the applications deployed on the website, to

accurately analyze how visitors navigate through the IRS website at an aggregated level. Data collected on a web user's (Taxpayer) session is depicted in section 1A of this PIA.

C. Employee – Only System Administrators, security personnel, and application support user ID and password are obtained directly from the employee user when they attempt to login to the system.

D. Other Third Party Sources – Contractor support will provide maintenance, support, and web hosting services for the IRS.GOV environment. However, requests for web content changes originate from the IRS (i.e., the IRS ultimately makes recommendations / change requests to update the content on the IRS.GOV website).

3. Is each data item required for the business purpose of the system? Explain.
All information supports the business processing and purpose of the system.

4. How will each data item be verified for accuracy, timeliness, and completeness?
The Content Management Application (CMA), which is an integrated part of the MITS–28 IRS.GOV GSS, provides reporting capabilities as a means to verify the accuracy and completeness of content on the website, as well as to monitor potential "broken links" as produced in an Orphan Content Report.

The Orphan Content Report is run on a monthly basis. If missing or incomplete fields exist, then the user is sent a request message asking for accurate and complete information.

5. Is there another source for the data? Explain how that source is or is not used.
Routine information comes directly from network traffic into the www.irs.gov infrastructure. No additional information on individuals, that has not previously been stated, is collected or stored. All individual user information will be collected on a completely separate, registered user portal (RUP) (i.e., not within the boundaries of this system or this PIA).

6. Generally, how will data be retrieved by the user?
Data / content retrieval by system users (backend data):

- System users must provide a user ID and password to access the system
- Depending on user role privileges, system users can retrieve report data on request.
- Actual web content and web pages (i.e., for web development / modification purposes) can be retrieved by the Content Provider (as defined in the CMA) via File Transfer Protocol (FTP) or Secure Shell (SSH) retrieval.

Data / content retrieval by taxpayers or web users (end user, web displayed data):

- Content displayed on web pages is accessed by going to the www.irs.gov public website.

7. Is the data retrievable by a personal identifier such as name, SSN, or other unique identifier?
Web user / taxpayer data is not retrievable by a personal identifier. System user actions are retrievable by User ID or Terminal ID from audit logs. Data that can be retrieved by those identifiers are the type and source of user events, as well as the time and date stamp the event occurred.

Access to the Data

8. Who will have access to the data in the system (Users, Managers, System Administrators, Developers, Others)?

Web users (Taxpayers, general public) will only have access to data displayed on the public website at www.irs.gov. System administrators, security personnel, and application support personnel will comprise of compilation of IRS employees, as well as contractor support personnel. Following are the categories of users that configure, operate, and maintain the IRS.GOV during normal use. Below is a list of users, both IRS employee and contractor support, that can have access to the system.

Role: IRS Users
Permission: IRS Content Management Application (CMA) User (IRS employees). The IRS CMA user will be an IRS employee who uses the CMA to create and/or manage the content of the IRS.GOV site.

Role: IRS Executives and IRS Employees
Permission: Web traffic reports, generated from the data collected by the SDCs without any unique identifiers (e.g., IP address), will be provided to various IRS executives and IRS employees through an IRS Intranet website.

Role: Contractor Users
Permission: Application Administrator (contractor support). Designated contractor IRS.GOV Technical Architecture Team members are responsible for maintaining the applications and software which reside in the IRS.GOV architecture. This includes database administrators, application administrators, and build managers.

Role: System [Site] Administrator (contractor support)
Permission: Designated IRS.GOV Site Administrators are responsible for the health and security of the site hardware and OS.

Role: Network Administrator (contractor support)
Permission: The Network Administrators manage switches, routers, and load balancing equipment. They have no physical access to other components, and only access systems for routine maintenance and troubleshooting.

Role: Managed Firewall Administrator (contractor support)
Permission: Firewall Administrators manage the network security perimeter of the infrastructure. They have no physical access to other components, and only access systems for routine maintenance and troubleshooting.

Role: Managed IDS Administrator (contractor support)
Permission: The Intrusion Detection Sensor (IDS) systems are managed remotely by a contractor team as part of their support to IRS.GOV.

Role: Help Desk (contractor support)
Permission: The Contractor Support Desk is allowed access to basic monitoring statistics of the www.irs.gov site. Help Desk Users coordinate with technical staff and have no accounts to login to the systems.

9. How is access to the data by a user determined and by whom?
All contractors are required to have a fully adjudicated background investigation prior to accessing IRS systems in keeping with their level of access to SBU and audit trail root data. In addition, the contracting officer's technical representative (COTR) maintains copies of all clearances. For IRS.gov, contractor staff have Contractor Moderate or High Risk background investigations. All users, both employee and contractor, receive access to the system only after going through the appropriate background check as stated in IRM 10.8.1 Information Technology (IT) Security Policy and Guidance.

Internal IRS user access is controlled by users completing the OL5081. OL5081 policies and procedures are followed whenever a new user needs to be granted access, an existing user needs to be granted additional access, or needs to be deleted from the system. All user rights are granted on a need to know basis; the least necessary access rights are granted. Additionally, access to the business applications is restricted by the menu–driven–approach, which only allows the authorized users to perform just the authorized functions.

All of the user roles stated in Item 8 of this document and again in the following list receive access in accordance with the requirements stated in IRM 10.8.1:

- IRS Content Management Application (CMA) User (IRS employees).
- Application Administrator (contractor support).
- System [Site] Administrator (contractor support).
- Network Administrator (contractor support).
- Managed Firewall Administrator (contractor support).
- Managed IDS Administrator (contractor support).
- Help Desk (contractor support).

10. Do other IRS systems provide, receive, or share data in the system? If YES, list the system(s) and describe which data is shared.
No. Although the Political Action Committee (PAC) 527 and Systemic Advocacy Management System (SAMS) reside on MITS–28, they do not send, receive, or share data with MITS 28. MITS 28 IRS.GOV is a general support system (GSS) that provides the underlying architecture and is not directly involved with the processing and sharing of data that is typically performed by Applications.

11. Have the IRS systems described in Item 10 received an approved Security Certification and Privacy Impact Assessment?
Not applicable.

12. Will other agencies provide, receive, or share data in any form with this system?
None.

Administrative Controls of Data

13. What are the procedures for eliminating the data at the end of the retention period?
System access records (for backend system users only) are retained in off–site storage for one year. IRS follows the disk sanitization procedures for destruction of discarded media. IRM 2.7.4, Management of Magnetic Media (Purging of Sensitive But Unclassified (SBU) Data and Destruction of Computer Media) provides those procedures used for sanitizing electronic media for reuse (e.g., overwriting or degaussing) and for controlled storage, handling, or destruction of spoiled media or media that cannot be effectively sanitized for reuse.

All IRS.GOV media is degaussed and reused or degaussed and destroyed when no longer useable. Paper documents are shredded and burned when disposed of. Excess hard drives are overwritten with an approved utility or degaussed using an approved degausser on the National Security Agency's (NSAs) Degausser Products List. Gdisk is currently used for the sanitization of IRS.GOV data. In addition, non–functioning tapes are appropriately degaussed and destroyed. Printouts and storage media are stored, handled, or destroyed in accordance with their levels of sensitivity and criticality. Users are provided security training on the proper handling, storage, and destruction of paper information. Locked shred bins may be located throughout the site to be used for disposing of sensitive and classified information. Local procedures are in place for proper destruction of outdated data diskettes and compact disks (CD).

The system administrator shall be responsible for conducting a quarterly review of all data storage media (disks, magnetic tapes, etc.) and their contents. Data media that is no longer required shall be returned for destruction. Individual files on multiple file disks shall be erased. The system administrator shall maintain a record of the dates the review was conducted as well as a current list of the media and its contents.

Specific to the IRS.GOV Contractor environment, contracted data retention facilities oversee the sanitization and disposal (i.e., destruction) of IRS.GOV media and stored records. Contractor management must authorize and approve the sanitization or disposal of all media. In addition, IRS.GOV Contractor environment media that contains program sensitive information is handled in a manner consistent with the handling of classified documents. As a result, sensitive media is properly marked and stored until sanitized or transferred.

14. Will this system use technology in a new way?
No. The web functionality, downloading ability (e.g., for IRS Forms and Publications), and calculators and tools are not a new use of technology.

15. Will this system be used to identify or locate individuals or groups? If so, describe the business purpose for this capability.
Yes, but only for special circumstances. In the event of an attack, this system could provide data, which when combined with other forensic information, may contribute to the capability to identify or locate an individual. However, it should be noted that the data in the system alone is not sufficient to locate an individual or group. Identification of an individual or group will only be possible when the data in the system is combined with other data, such as access log information provided by an Internet Service Provider (ISP). This would be an official law enforcement action accomplished via a system external to the IRS.gov website, and therefore outside the purview of this PIA.

16. Will this system provide the capability to monitor individuals or groups? If yes, describe the business purpose for this capability and the controls established to prevent unauthorized monitoring.
No.

17. Can use of the system allow IRS to treat taxpayers, employees, or others, differently?
No. Disparate treatment of individuals or web users is not intended. However, in the event of an attack on the www.irs.gov website (e.g., hacker infiltrating the system through the web) or infrastructure, individual IP addresses may be barred from accessing the system at the time of the attack as a method to protect the system.

18. Does the system ensure "due process" by allowing affected parties to respond to any negative determination, prior to final action?
Not applicable.

19. If the system is web–based, does it use persistent cookies or other tracking devices to identify web visitors?
The MITS 28 IRS.GOV GSS supports the www.irs.gov website and implements persistent cookies based on the White House Open Government initiative, OMB Memorandum 10–22. Persistent cookies enhance web metrics and helps the IRS analyze how visitors navigate through www.irs.gov. Persistent cookies will allow the IRS the capability to analyze the effectiveness of the website, make design improvements to enhance site performance, which will improve customer satisfaction. We do not collect personally identifiable information (PII), about visitors to our website.

The IRS uses ”temporary” or “session” cookies on some of the applications deployed on our website, as well. Session cookies allow the IRS to accurately analyze how visitors navigate through the IRS website at an aggregated level. Session cookies also allow the IRS to perform site improvements based upon the way the website is actually used by the visitors. These cookies are stored in memory and are only available during an active web browser session. Visitors can adjust their web browser to accept or decline web cookies, or to alert them when cookies are in use.

In the CMA, the internal IRS user has the option of using a persistent cookie to store their CMA User ID. This allows their User ID to be “auto–filled” in their browser, so that the IRS CMA user only types their password to login. This enables customization to accommodate personal preferences and is voluntarily enabled by the user.

View other PIAs on IRS.gov