Department of the Interior
Privacy Impact Assessment

Name of Project: Enterprise Web
Bureau: U.S. Geological Survey
Project's Unique ID: 010-12-03-00-01-3004-00

A. CONTACT INFORMATION:

Nancy Sternberg
MS 159 12201 Sunrise Valley Drive
Reston, VA 20192
Phone: 703-648-6861
nsternberg@usgs.gov

B. SYSTEM APPLICATION/GENERAL INFORMATION:

1) Does this system contain any information about individuals?

Yes, business contact information about USGS employees and persons representing other governmental agencies or organizations, including name, business address, business phone number and business email address is available from Enterprise Web hosted site pages. Additionally, a hosted web application allows members of the public to input their email addresses or phone numbers to participate in the USGS Threshold Alert Notification Service.

a. Is this information identifiable to the individual[1]? Yes

b. Is the information about individual members of the public?

Yes

c. Is the information about employees? Yes

2) What is the purpose of the system/application?

[1] "Identifiable Form" -According to the OMB Memo M-03-22, this means information in an IT system or online collection: (i) that directly identifies an individual (e.g., name, address, social security number or other identifying number or code, telephone number, email address, etc.) or (ii) by which an agency intends to identify specific individuals in conjunction with other data elements, i.e., indirect identification. (These data elements may include a combination of gender, race, birth date, geographic indicator, and other descriptors).

The purpose of the Enterprise Web IT System is to provide robust, secure and fault tolerant web hosting services to USGS programs for their delivery of USGS science. The Enterprise Web system serves as a Web virtual-hosting provider available for use to all USGS Web site administrators. This virtual hosting capability (http://natweb.usgs.gov/ (internal USGS access only) provides seamless public access to locally managed data and information available using the model of a distributed system of mirrored servers located at USGS high bandwidth locations (Menlo Park, Reston, and Sioux Falls centers) in proximity of the boundary routers. The systems of servers at high bandwidth locations are also known as modules and provide for high reliability and load sharing of user requests. Additionally, Enterprise Web hosts backend content and applications for the USGS homepage suite at www.usgs.gov and USGS multimedia gallery at gallery.usgs.gov.

3) What legal authority authorizes the purchase or development of this system/application?

The U.S. Congress provides financial support of Enterprise Web through appropriations to the USGS.

C. DATA in the SYSTEM:

1) What categories of individuals are covered in the system?

Enterprise Web hosted sites may contain business contact information about individuals who are i) USGS employees, or ii) are representing other governmental agencies or organizations as collaborators with USGS employees, iii) are members of USGS steering committees, coordination groups, working groups or subcommittees, or iv) members of the general public.

2) What are the sources of the information in the system?

a. Is the source of the information from the individual or is it taken from another source? If not directly from the individual, then what other source?

Information about USGS employees is obtained from the USGS staff directories which have PIAs. Information about persons representing other governmental agencies, organizations or the public is provided by the individual.

b. What Federal agencies are providing data for use in the system?

None, data are only from USGS.

c. What Tribal, State and local agencies are providing data for use in the system?

None.

d. From what other third party sources will data be collected?

None.

c. What information will be collected from the employee and the public?

Enterprise Web hosted sites do not collect information from employees. Information collected from the public will include only e-mail addresses and/or telephone numbers. Telephone numbers are only collected if the end user requests alerts be sent to their cell phones in the form of text messages.

3) Accuracy, Timeliness, and Reliability

a. How will data collected from sources other than DOl records be verified for accuracy?

Enterprise Web hosted sites depend on the individuals providing their business information to verify their data for accuracy. All information collected from the public is input by the end user. It must also be verified by the end user before automated messages are sent from the web application.

b. How will data be checked for completeness?

As employee information is obtained from either USGS Lotus Notes System, Active Directory, or USGS Geospatial Information Management System, those systems are relied on the check for completeness. Enterprise Web hosted sites depend on the individuals providing their business information to verify their data for completeness. Members of the public will verify their information via e-mail/web before they receive automated messages.

c. Is the data current? What steps or procedures are taken to ensure the data is current and not out-of-date? Name the document (e.g., data models).

The employee's information from the phonebook is current as of the time the information was obtained from either the USGS Lotus Notes System, Active Directory, or the USGS Geospatial Information Management System (GMIS). Information about individuals representing other governmental agencies or organizations is updated when the site staff are notified of changes. Information about the public is current as of their action to subscribe to the USGS Threshold Alert Notification Service.

d. Are the data elements described in detail and documented? If yes, what is the name of the document?

Not applicable, Enterprise Web hosted sites do not collect information from

employees or the public from the site; information is only provided.

D. ATTRIBUTES OF THE DATA:

1) Is the use of the data both relevant and necessary to the purpose for which the system is being designed?

Yes.

2) Will the system derive new data or create previously unavailable data about an individual through aggregation from the information collected, and how will this be maintained and filed?

No.

3) Will the new data be placed in the individual's

record? No.

4) Can the system make determinations about employees/public that would not be possible without the new data?

No.

5) How will the new data be verified for relevance and accuracy?

New data provided by members of the public for purposes of alerts will be verified by the end user, only.

6) If the data is being consolidated, what controls are in place to protect the data from unauthorized access or use?

Not applicable, data are not consolidated.

7) If processes are being consolidated, are the proper controls remaining in place to protect the data and prevent unauthorized access? Explain.

Not applicable, no processed consolidated.

8) How will the data be retrieved? Does a personal identifier retrieve the data? If yes, explain and list the identifiers that will be used to retrieve information on the individual.

USGS employee business contact information is retrieved from the Lotus Notes system profile, Active Directory, and by USGS Geospatial Information Management System (GMIS) using the person's name. Data provided by the members of the public for

USGS Threshold Alert Notification Service are stored in tab-delimited files and are retrieved by web applications in order to route alerts.

9) What kinds of reports can be produced on individuals? What will be the use of these reports? Who will have access to them?

Not applicable, no reports are produced.

10) What opportunities do individuals have to decline to provide information (i.e., where providing information is voluntary) or to consent to particular uses of the information (other than required or authorized uses), and how individuals can grant consent.)

Not applicable, providing any information to Enterprise Web hosted sites is optional.

E. MAINTENANCE AND ADMINISTRATIVE CONTROLS:

1) If the system is operated in more than one site, how will consistent use of the system and data are maintained in all sites?

The Enterprise Web system operates from 3 modules located in Reston, Virginia; Sioux Falls, South Dakota; and Menlo Park, California. Enterprise Web's core infrastructure model processing flow consists of three (3) modules designed to replicate Web data and services so that information passed to the public via WWW is secure, highly available, and load balanced. Access to update Web content contained within the Enterprise Web modules occurs using the AFS system clients which then update files provided that valid AFS's Kerberos authentication has occurred. The AFS servers work in Master/Slave relationships with units of storage called volumes. When a file is updated, the master AFS server for the volume containing that file accepts the change. When a volume is "released," all changes on that volume are mirrored to all AFS servers defined for that volume as read-only copies.

2) What are the retention periods of data in this system?

The retention periods, or time periods for which backup copies are kept as well as how many versions, or cycles, of the same backups are established in accordance with DOI Records Management Policy and more specifically the DOl Information Technology Security Policy Handbook, Section CP-9, Information System Backup:

Bureaus and offices shall conduct backups of user-level and system-level information (including system state information) contained in all information systems at least weekly and protect backup information at the storage location.

Specifics include:

Nightly full backups of each AFS volume are stored server hard disks. On the NatWeb web log analysis are a triplet of root cron tasks.

This task in turn runs a program against each of the three AFS virtual servers in the given region.

Disk images of selected AFS volumes are dumped to the local disks.

Nightly backups of the server hard disks are stored onto tape backup.

Additionally, public data are kept as long as the end-user allows. End-users can automatically remove their information from the USGS Threshold Alert Notification Service without webmaster intervention.

3) What are the procedures for disposition of the data at the end of the retention period? How long will the reports produced be kept? Where are the procedures documented?

Disposition procedures for this data will be in accordance with DOI Media Sanitation Policy and Procedure as described in Section MP-6 of the DOI Information Technology Security Policy Handbook. In general, tapes are overwritten and reused. If the tape does not function, the Backup Administrator is notified and the tape is sent to the degaussing unit who destroys the tape.

Members of the public can delete information without webmaster intervention. These procedures are documented both in e-mail.

4) Is the system using technologies in ways that the DOl has not previously employed (e.g., monitoring software, Smart Cards, Caller-ID)?

No.

5) How does the use of this technology affect public/employee privacy?

Not applicable. No Privacy or Sensitive data are stored or processed in this application.

6) Will this system provide the capability to identify, locate, and monitor individuals? If yes, explain.

No. This system does not provide the capability to identify, locate or monitor individuals. The only location information provided is the individual's business contact information such as work location, phone number or email address.

7) What kinds of information are collected as a function of the monitoring of individuals?

Not applicable, information is not collected to allow the monitoring of individuals.

8) What controls will be used to prevent unauthorized

monitoring? Not applicable.

9) Under which Privacy Act systems of records notice does the system operate? Provide number and name.

Department of Interior Social Networks, DOI-08

10) If the system is being modified, will the Privacy Act system of records notice require amendment or revision? Explain.

Not applicable.

F. ACCESS TO DATA:

1) Who will have access to the data in the system? (E.g., contractors, users, managers, system administrators, developers, tribes, other)

There are no user login accounts on the Enterprise Web systems, only administrators have login access.

Regarding email addresses and phone numbers provided by members of the public, Public users will always have access to delete their entries. Only the webmaster will have access to view the list of e-mail addresses.

2) How is access to the data by a user determined? Are criteria, procedures, controls, and responsibilities regarding access documented?

Users are given accounts to use the system only. This allows them to access only the content they maintain. System administrators are given UNIX system level accounts.

3) Will users have access to all data on the system or will the user's access be restricted? Explain.

Users can only access content in Enterprise Web through their AFS client. Their AFS Login allows them access to only the content they maintain.

4) What controls are in place to prevent the misuse (e.g., unauthorized browsing) of data by those having access? (Please list processes and training materials)

All users must agree to the USGS Information Technology Rules of Behavior before access is granted to any USGS system, including an AFS account through Enterprise Web. Additionally, logging and auditing are enabled on the system. All logs are retained (online) for a period of 12 months, after which they are deleted from the system. Specific audit requirements including logon/logoff, program initiation, deletion of objects, changes to user rights, admin/root access, changes to OS and access to security files are being met through the use of both the logcheck

utility (described at) and the open source software Integrit and Big Brother. Logcheck scans system log files every 15 minutes on the servers, and emails system administrators of any unusual events found, including logon/logoff and admin/root access. Integrit creates a checksum DB of system files every 6 hours, and emails system administrators of changes to the DB caused, for example, by deletion of objects, changes to OS files or access to security files. Integrit reports are also monitored by Big Brother in the "ts" column at.

Additionally, public data collected including email address and phone numbers are stored in flat data file on the AFS system. Permissions are set on the data file such that only the webmaster and the web daemon have access to the file.

5) Are contractors involved with the design and development of the system and will they be involved with the maintenance of the system? If yes, were Privacy Act contract clauses inserted in their contracts and other regulatory measures addressed?

Contractors are involved with development of the system. A Privacy Act contract clause is inserted into their GSA contracts.

6) Do other systems share data or have access to the data in the system? If yes, explain.

No.

7) Who will be responsible for protecting the privacy rights of the public and employees affected by the interface?

Not applicable, no privacy act data.

8) Will other agencies share data or have access to the data in this system (Federal, State, and Local, Other (e.g., Tribal))?

No.

9) How will the data be used by the other gency?

Not applicable.

10) Who is responsible for assuring proper use of the data?
The System Manager