National Science Foundation

Research.gov
Privacy Impact Assessment
Version: 1.1

Original Date: 12/05/07(Updated 8/20/08)

Table of Contents

Revisions

Revision Number	Author	Date	Description
1.0	Research.gov Team	Dec 12, 2007	Original version
1.1	M Tillotson	Aug 20, 2008	Updated 4.4 Maintenance of Admin Controls question 11 with NSF 72 Research.gov privacy notice

1. BACKGROUND

The Privacy Impact Assessment (PIA) is a vehicle to address privacy issues in information systems. The PIA template establishes requirements for addressing privacy during the information systems life cycle.

1.1 Organizational Background

The Division of Information Systems within the Office of Information Resources Management is leading the overall implementation of Research.gov. Research.gov is a web based portal that will serve the National Science Foundation and agencies who partner with NSF in Research.gov, and the larger research community.

2. SCOPE

Protecting an individual's right to privacy is predicated on various Federal laws, directives, and standards; the overarching Federal laws being the Privacy Act of 1974 and the more recent E-Government Act of 2002. Federal guidance requires that, where possible, the PIA process be integrated into the General Support Systems (GSS)/Major Applications (MA) life cycle.

3. ENVIRONMENT

Research.gov Overview

Research.gov is a portal for information and not a new Privacy Act collection. Research.gov is a research-oriented internet web portal solution to deliver Grants Management Line of Business (GMLoB) consortium services. Led by NSF, Research.gov improves customer service for applicants and awardees by streamlining and standardizing grant business processes among partner agencies. The Department of Agriculture's Cooperative State Research, Education, and Extension Service (CSREES) is NSF's partner agency. The Research.gov web portal will provide a comprehensive menu of grants management services for both research institutions and partner agencies.

The initial release scheduled for December 2007 will be available for all principal investigators and sponsored project office officials to view grant application status for NSF and CSREES. Research.gov will also display publicly available information including research highlights, a policy library, public award information, and publication citations. Additional services will be developed and deployed over time in conjunction with additional partner agencies. The Research.gov PIA will be updated as new services and functions are added and other agencies join.

4. PRIVACY IMPACT ASSESSMENT CRITERIA

The following sections contain the appropriate questions that are used to collect the required GSS/MA information. The NSF Privacy Officer and other reviewing officials will review the results to ensure that an individual's personal identifiable information is adequately secure.

4.1 Data in the System

The sources of the system information are an important privacy consideration. The information becomes especially important if the data is gathered from sources other than NSF records. Information collected from non-NSF sources should be verified, to the extent practicable, for accuracy, that the information is current, and the information is complete. Accurate information is important if the information will be used to make determinations about individuals.

Privacy Criteria	Descriptive Response
1. Provide a general description of the information type (i.e., persons name, SSN, etc.) to be collected or processed by the GSS or MA.	Research.gov is not a new collection of information. Research.gov is an internet web portal that displays information previously collected by agency grant management systems (for NSF, FastLane, for CSREES, C-REEMS).
2. What are the sources of the information in the system? (Note: This is an important privacy consideration if the data is gathered from other than NSF records).	Research.gov will display information that was previously provided by authorized external users (i.e., scientists, educators, technology experts, research administrators, graduate students, and panelists) over the Internet. Research.gov will: • read Application Status data from FastLane and USDA/CSREES's staging table • read Award data from NSF's FastLane, Proposal and Reviewer System (PARS), Awards, Institution, Financial and Accounting System, the Central Contractor Registration (CCR) system, and USDA/CSREES's staging table • read/write Institution information from FastLane, the Central Contractor Registration (CCR) system, and Research.gov's Access Manager repository, as well as USDA/CSREES's staging table (basic institutional profile information as well as employee roles and access privileges)
3. What NSF files and databases are used?	Data from the NSF proposal and awards databases is used.
4. What other Federal Agencies, if any, are providing data for use in the system?	The following agencies currently provide data for Research.gov: USDA/CSREES (Partner Agency), DoD/CCR (institution data provider), USDA (E-

Privacy Criteria	Descriptive Response
	Authentication Credential Service Provider (CSP)), and Operational Research Consultants (ORC). Any future Research.gov partner agencies and/or E-Authentication CSPs will also provide data for use by the system.
5. From what other third party sources will data be collected?	N/A
6. What information will be collected from the employee?	Research.gov does not collect information. Research.gov is a portal that displays information previously collected through FastLane and USDA under NSF Privacy Act Systems 12, 50, 51 and USDA/CSREES-4 Report.
7. If data is collected from sources other than NSF records, how is it being verified for accuracy? *(Note: This is especially important if the information will be used to make determinations about individuals).*	USDA/CSREES (Partner Agency) and any other future partner agencies: CSREES is responsible for their data validation. NSF maintains configuration control over the schema of CSREES' staging table. DoD/CCR (institution data provider): CCR is being used and trusted as the authoritative source for DUNS numbers. This data is used by other e-gov initiatives. USDA and Operational Research Consultants (ORC) (E-Authentication CSP): The USDA and ORC CSPs are being used as a trusted source for valid e-authentication credentials.
8. How will data be checked for completeness?	These business rules include the validation of the data's completeness via the existing FastLane User Registration process.
9. Is the data current? How do you know? What mechanisms were used to validate the data's currency?	Research.gov is an internet web portal that displays information previously collected by agency grant management systems (for NSF, FastLane, for CSREES, C-REEMS). Users of those source systems validate the accuracy of the data.
10. What data elements are described? What level of detail is used in documenting data elements?	The Research.gov data dictionary contains a list of all data fields used by the system, including data name, data type (i.e., alpha numeric or text), and data length.
11. If data elements are documented, what is the name of the document?	Research.gov Data Dictionary

4.2 Access to the Data

Who has access to the data in a system must be defined and documented. Users of the data can be individuals, other systems, and other agencies. Individuals who have access to the data can be system users, system administrators, system owners, managers, and developers. When individuals are granted access to a system, their access should be limited, where possible, to only that data needed to perform their assigned duties. If individuals are granted access to all of the data in a system, procedures need to be in place to deter and detect browsing and unauthorized access. Other systems are any programs or projects that interface with the system and have access to the data.

Privacy Criteria	Descriptive Response
1. Who has access to the data in the system? (Note: Users of the data can be individuals, other systems, programs, projects, or other agencies. Individuals who have access to the data can be system users, system administrators, system owners, managers, and developers).	Research.gov has role based access. The roles are: • Institution Administrator – Add, Modify, Users, Change User Permissions • Sponsored Project Officer (SPO) – View proposal status • Principal Investigator (PI)/Project Director (PD) – View their own proposals, modify their contact information • Content Management Editor – NSF and Partner agency staff who edits and creates content • Content Management Approver – NSF agency staff who approves content External users are granted access to the system by their SPOs. For Research.gov, the users are assigned various roles – Principal Investigator, Sponsored Projects Officer or Institutional Administrator. In addition, NSF staff work with various organizations to establish the initial ability to access Research.gov. When an institution registers with Research.gov, the institution is required to obtain the signature of an Authorized Organizational Representative (AOR). The AOR and other institutional personnel, as designated by AOR, agree to accept the responsibility for protection of the data in Research.gov system. Other individuals with access include the Research.gov system administrators, database administrators, and members of the NSF operational support teams.

Privacy Criteria	Descriptive Response
2. Where individuals are granted access to all of the data in a system, what procedures are in place to deter and detect browsing and unauthorized access?	All requests for access to data in the system are validated through system access controls to ensure the user is both authenticated and authorized to access that information. Users are granted access to only those functions required to complete their job responsibilities. Institution Administrators, Sponsored Project Officers, and Principal Investigators have specialized access to restricted Research.gov functions. Only the Research.gov database administrators have access to all of the data in the Research.gov database. Such access is necessary to develop and maintain the system, as well as provide customer support.
3. When individuals are granted access to a system, how is their access being limited, where possible, to only that data needed to perform their assigned duties?	The Research.gov authorization model defines specific roles and uses role-based access control to ensure that an individual's access is consistent with the need to complete his or her assigned duties. Users are granted access to the system by an Institution Administrator in the same manner that their access is granted in FastLane. For system owners, administrators, or developers, the level of access is controlled by the NSF system administration policy.
4. How or what tools are used to determine a user's data access?	Institutional accounts and roles are provisioned by an authorized administrator within the institution, using an interface provided by Research.gov which builds on the current FastLane account registration and management modules. For system owners, administrators, or developers, accounts are controlled by system administration staff.
5. Describe the criteria, the procedures, the controls, and the responsibilities in place regarding the manner in which data access is documented.	Organizations cannot use Research.gov until they register with NSF and the registration is accepted. Research.gov data access is documented in the registration form that is completed online; signed by the organization's Authorized Organizational Representative; faxed to NSF; reviewed by NSF; and then accepted or rejected and stored.
6. Do other systems share data or have access to data in this system? If yes, explain.	The Research.gov institution and user registration data is stored in the FastLane database. The management of Research.gov users and storage of their authorization data is isolated from other

Privacy Criteria	Descriptive Response
	applications. The USDA/CSREES staging tables and USDA E-Authentication also share data with Research.gov during various stages of the login process or during use of Research.gov services.
7. Who has the responsibility for protecting the privacy rights of the individuals affected by any system interface?	The Research.gov Project Management Office (PMO) is responsible for protecting the rights of the individuals affected by any system interface. The PMO coordinates very closely with the NSF Office of the General Counsel and its partner agencies on all privacy issues.
8. Will other agencies share data or have access to data in this system?	No.
9. How will the NSF use this data?	To provide and administer grants management services for the customers of Research.gov.
10. Who is responsible for assuring proper use of the data?	External users (i.e., scientists, educators, technology experts and administrators); internal NSF operational teams; and NSF users ensure proper use of Research.gov data.
11. How will the system ensure that agencies only get the information they are entitled to?	Other agencies do not have direct access to the system.

4.3 Attributes of the Data

When requirements for the data to be used in the system are being determined, those requirements must include the privacy attributes of the data. The privacy attributes are derived from the legal requirements imposed by the Privacy Act of 1974. First, the data must be *relevant and necessary* to accomplish the purpose of the system. Second, the data must be *complete, accurate and timely*. It is important to ensure the data has these privacy attributes in order to assure fairness to the individual in making decisions based on the data.

Privacy Criteria	Descriptive Response
1. Explain how the use of the data is both relevant and necessary to the purpose for which the system is being designed?	Research.gov data supports NSF and its partner agencies' grants management lifecycle.
2. Will the system derive new data or create previously unavailable data about an individual through aggregation for the	Research.gov does not derive or create new data.

Privacy Criteria	Descriptive Response
information collected?	
3. Will the new data be placed in the individual's record?	N/A
4. Can the system make determinations that would not be possible without the new data?	N/A
5. How will the new data be verified for relevance and accuracy?	N/A
6. If data is being consolidated, what controls are in place to protect the data from unauthorized access or use?	Research.gov does not consolidate or aggregate data.
7. If processes are being consolidated, are the proper controls remaining in place to protect the data and prevent unauthorized access? Explain	N/A
8. How will the data be retrieved? Can the data be retrieved using a personal identifier (i.e., name, address, etc.)? If yes, explain.	Individual users can retrieve their own data through the registration, log in, and authorization procedures described above. Sponsored Projects Officers are able to see applications for their institution/office based on the institution's unique identifier (DUNS ID) they provide.
9. What are the potential effects on the due process rights of individuals with respect to the following: • Consolidation and linkage of files and systems; • Derivation of data; • Accelerated information processing and decision-making; • Use *of new technologies?*	No effect.
10. How will these effects be mitigated?	There are no effects.

4.4 Maintenance of Administrative Controls

Automation of systems can lead to the consolidation of processes, data, and the controls in place to protect the data. When administrative controls are consolidated, they should be evaluated so that all necessary controls remain in place to the degree necessary to continue to control access to and use of the data.

Data retention procedures should be documented. Data retention procedures require review to ensure they meet statutory requirements. Rules must be established for the length of time information is kept and for assuring that it is properly eliminated (i.e., archived, deleted, etc.) at the end of that time.

The intended and potential monitoring capabilities of a system must be defined and safeguards must be installed to ensure privacy and prevent unnecessary intrusion.

Privacy Criteria	Descriptive Response
1. Explain how the system and its use will ensure equitable treatment of individuals.	N/A.
2. If the system is operated in more than one site, how will consistent use of the system and data be maintained in all sites?	Research.gov accesses data located at partner agency locations, and its users are geographically distributed. NSF, as the Research.gov lead, ensures consistent use of the system and data through administrative controls such as Memoranda of Understanding and Service Level Agreements.
3. Explain any possibilities of disparate treatment of individuals or groups.	Research.gov processing does not create possibilities for disparate treatment of individuals or groups.
4. What are the retention periods of data in this system?	N/A
5. What are the procedures for eliminating the data at the end of the retention period? Where are the procedures documented?	N/A
6. While the data is retained in the system, what are the requirements for determining if the data is still sufficiently accurate, relevant, timely, and complete to ensure fairness in making determinations?	N/A
7. Is the system using technologies in ways that NSF has not previously employed? How does the use of this technology affect individual's privacy?	This technology and communications to partner agencies has not previously been employed by NSF. Research.gov is using system-to-system interfaces to provide grants management services on behalf of its agency partners. The use of this new technology does not affect an individual's privacy.
8. Will this system provide the capability to identify, locate, and monitor individuals? If yes, explain.	Research.gov monitors user access and logs certain user transactions to ensure system performance and provide auditing capabilities (see Research.gov Web Privacy Policy).

Privacy Criteria	Descriptive Response
9. Will this system provide the capability to identify, locate and monitor groups of people? If yes explain.	No.
10. What controls will be used to prevent unauthorized monitoring?	All requests for access to data in the system are validated through system access controls to ensure the user is both authenticated and authorized to access that information. Users are granted access to only those functions required to complete their job responsibilities. Institution Administrators, Sponsored Project Officers, and Principal Investigators have specialized access to restricted Research.gov functions. Only the Research.gov database administrators have access to all of the data in the Research.gov database. Such access is necessary to develop and maintain the system, as well as provide customer support.
11. Under which System of Record notice does the system operate? Provide number and name.	NSF 72 Research.gov
12. If the system is being modified, will the System of Record require amendment or revision? Explain	N/A

Additional Assistance

For additional assistance with completing this assessment, you may contact NSF Privacy Act Officer, Leslie Jensen, at 703 292 5065 or the NSF Privacy Advocate, Mary Lou Tillotson, at 703 292 4264.

Review

When the PIA is complete, please submit the PIA to the NSF Privacy Act Officer for review at ljensen@nsf.gov and to the NSF Privacy Advocate at mtillots@nsf.gov.