National Science Foundation

Privacy Impact Assessment

for the

Scientists and Engineers Statistical System (SESTAT)

November 2022

1.   CONTACT INFORMATION

System Owner
Tiffany Julian
NCSES
(703) 292-5159
tjulian@nsf.gov

2.   GENERAL SYSTEM INFORMATION

1. **Name of system:**

Scientists and Engineers Statistical Data System (SESTAT) http://www.nsf.gov/statistics/sestat/

2. **Description of system or electronic collection of information and its purpose.**

SESTAT is an integrated data system providing information about the employment, educational, and demographic characteristics of the college-educated science and engineering workforce in the United States. SESTAT does not collect or process any personally identifiable information (PII). Statistical data is provided by SESTAT in the following formats:

- Surveys and other quantitative information about people in the United States with education and a profession in science, engineering, or technology

- Detailed Statistical Tables (DSTs) containing extensive tabulated survey data

- Publications from surveys and analyses

- Overview publications issued on a periodic basis

- Special publications issued on an as-needed basis

3. **What is the purpose of the system or electronic collection of information?**

SESTAT provides a web-based central clearinghouse for the collection and analysis of statistical data regarding the science and engineering workforce. Data is collected by NSF Social Behavior and Economic Sciences (SBE), National Center for Science and Engineering Statistics (NCSES).

4. **Requested Operational Date?**

SESTAT has been operational since 1995.

5. **Does the collection create a new Privacy Act System of Records Notice (SORN), or is the PII collection covered by one or more existing SORNs? If so, name the SORN.**

PII collected and maintained by SESTAT is covered by the following Privacy Act system of records:

- NSF 43, "Doctorate Work History File"
- NSF 58, "National Survey of Recent College graduates and Follow-up File"

6. **What specific legal authorities, arrangements and/or agreements require collection?**

Authority for the collection of information from scientists and engineers is provided by 42 U.S.C. 1862 (a)(3), 1862 (c)(6), 1863 (j)(1), 1863 (j)(2), and 1885 (d).

3.   PII IN THE SYSTEM

1. **What PII is to be collected, used, disseminated or maintained in the system or collection?**

SESTAT does not collect or process PII. SESTAT contains aggregated data that includes the following:

- Demographic Information – Age, race, ethnicity, sex, disability, country of birth, marital status, number and age of children living at home, citizenship, and geographic location.

- Educational Information – Academic degrees earned, major subject of degree(s), earlier education, continuing education, and parent's educational attainment.

- Employment Information – Employment status, field of employment, type of employer, professional activities, labor force status, salary, and spouse's employment status.

2. **What are the sources of the PII?**

Data collected by SESTAT is provided by individuals holding science and engineering degrees, through the following voluntary surveys:

- National Survey of College Graduates (NSCG) - conducted for National Science Foundation by U.S. Department of Commerce, Economics and Statistics Administration, US Census Bureau.

- National Survey of Recent College Graduates (NSRCG) - Mathematica Policy Research, Inc. conducted the first stage of data collection for SRS under contract in 2006. U.S. Census Bureau conducted the second stage of data collection.

- Survey of Doctorate Recipients (SDR) - conducted for NSF by National Opinion Research Center at University of Chicago.

3. **What technologies will be used to collect the PII?**

SESTAT data is collected using mail survey forms, web survey forms, and computer assisted telephone interviews.

4.   ATTRIBUTES OF THE DATA (USE AND ACCURACY)

1. **Describe the uses of the PII.**

Data is aggregated for use in analyzing current trends in science and technology education and research.

2. **Does the system perform any strictly analytical functions on the PII?**

   Data is analyzed by SBE/NCSES personnel and is presented in publications as summary survey findings. Aggregated data can also be analyzed by SESTAT users through the SESTAT Data Tool.

3. **How will the accuracy of the PII collected from individuals or derived by the system be ensured?**

   Organizations such as the US Census Bureau, Mathematica Policy Research, and National Opinion Research Center of the University of Chicago conduct data collection for NSF SESTAT. NSF contract specifications include provisions that assure data accuracy. There are three phases of survey data collection, and each phase has mechanisms to assure data accuracy as follows:

   a. The initial phase is data collection via paper questionnaire, phone survey, or web survey. Organizations have contracts with NSF SESTAT to perform data collection and are bound to contract specifications.

   b. The second phase is data integration where inaccurate and irrelevant parts of the data files are identified and then replaced, modified, or deleted as required. Organizations have contracts with NSF SESTAT to perform data integration and are bound to contract specifications.

   c. The third phase is quality checking where data is validated. Individuals or organizations have contracts with NSF SESTAT to perform quality checking and are bound to contract specifications.

5.   SHARING PRACTICES

1. **Describe any sharing of the PII with internal or external organizations.**

   - Public-use data is available through the NSF website.

   - Licensees use SESTAT data as a resource for summary statistical information about science education and professions.

   - U.S. Federal and state government agencies use SESTAT data for developing policy and funding programs related to scientific and technical education and research.

2. **How is the PII transmitted or disclosed to the internal or external organization?**

   SESTAT users can retrieve public-use data via the SESTAT home page using a link to the SESTAT Data Tool; links to completed publications; and a link for submitting a download request for a SESTAT public-use data file.

3. **How is the shared PII secured by external recipients?**

   Licensees for restricted-use microdata files are bound by the terms of the license, which specifies stringent data privacy and security requirements that must be met and maintained. The restricted use data Licensing Guide is at http://www.nsf.gov/statistics/license/.

6.   NOTICE TO INDIVIDUALS TO DECLINE/CONSENT USE

1. **How does the program or collection provide individuals notice prior to the collection of information? If notice is not provided, explain why not.**

   SESTAT questionnaires are voluntary and include a notification that states the specific use(s) of the data collection. Survey questionnaires may be found at the following link: http://www.nsf.gov/statistics/question.cfm#ScienceandEngineeringWorkforce.

2. **Do individuals have the opportunity and/or right to decline to provide any or all PII?**

   The surveys are voluntary. Individuals provide only the information that they choose. The surveys notify individuals that "failure to provide some or all of the requested information will not in any way adversely affect you."

3. **Do individuals have the right to consent to uses of their PII?**

   When choosing to provide survey information, an individual consents to the uses specified in the survey notification.

7.   ACCESS TO DATA (ADMINISTRATIVE AND TECHNICAL CONTROLS)

1. **What categories of individuals will have lawful access to the system?**

   Public Users – Public users have access to only public-use data.

   Restricted Users – Users that have access to non-aggregated microdata files. The microdata files do not contain PII.

   - NSF staff members have defined classes or sets of access privileges to SESTAT that are commensurate with the level of access required to perform their official duties.

   - Licensees are granted tightly regulated access to microdata files through the NCSES licensing process.

2. **How is permissible access by a user determined? Are procedures documented?**

   - Each class of SESTAT users has a defined set of privileges and data access.

   - Public users do not need to authenticate to SESTAT using a user ID and password, and do not sign Terms and Conditions statements. Public users have access to data that is not restricted.

   - Restricted users are required to authenticate using a user ID and password. Restricted users are required to accept NSF warning notices prior to authentication. Restricted users are limited to privileges and data in each assigned class.

3. **What auditing measures/controls and technical safeguards are in place to prevent exposure or misuse of PII by authorized users, e.g., record browsing, extraction?**

Specific software events are audited that document the access and use of SESTAT. The events are recorded in system logs to permit the detection and/or prevention of unauthorized access or inappropriate usage. The logging of a specific event may be turned on or off, or otherwise adjusted, depending on revised threat assessments or system performance and cost considerations.

4. **Describe privacy training provided users, general or specific, relevant to the program or system?**

- SESTAT web pages provide access to the NSF Privacy Policy via a Privacy Policy link.

- NSF staff receive training on the Privacy Act through annual Security and Privacy Awareness Training. NSF Security and Privacy Awareness Training provides the protection requirements for NSF sensitive information; the various federal laws and guidance that relate to the protection of privacy in individuals and business; and an introduction to NSF's privacy policies.

- The SESTAT Data Tool has a banner explaining the Terms and Conditions of use.

5. **Describe the extent to which contractors will have access to the system.**

Contractors may be assigned to a SESTAT user class and are limited to the permissions of the assigned class. NSF employees as well as NSF contractors are trained in privacy principles.

6. **Describe the retention period for the personal records in the system.**

Data in SESTAT is retained indefinitely.

7. **What is the disposition of the personal records at the end of the retention period?**

N/A

8.   SECURITY

Is the PII secured in accordance with FISMA requirements?

NSF operates SESTAT in accordance with security procedures required by federal law and policy to ensure that information is appropriately secured. NSF has conducted a risk assessment, identified appropriate security controls to protect against identified risk, and implemented those controls. NSF performs monitoring, testing, and evaluation of controls on a regular basis to ensure controls continue to work properly.

9.   PRIVACY ANALYSIS

The risks of divulging private information (contact information as listed in 3.1.) displayed are mitigated by:

- User roles determine what information and records a specific internal NSF user can view or edit.
- Database administrators have access to the database and are trained to know what is considered proper access.
- All NSF staff and contractors are required to annually complete IT Security and Privacy Awareness training.