NATIONAL SCIENCE FOUNDATION

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION
NATIONAL SCIENCE FOUNDATION
EDUCATION AND TRAINING APPLICATION

Executive summary

This request is for the renewal of OMB approval for the U.S. National Science Foundation's "Education and Training Application." This data collection is critical for monitoring, research, and evaluating NSF's investments in STEM education and workforce development, and essential to continue fulfilling statutory responsibilities under the *Foundations for Evidence-Based Policymaking Act of 2018* (Public Law 115-435).

The proposed changes to the currently approved collections are outlined in Section A.15. Most of these changes reflect the system's evolution from a pilot phase to an implementation phase, along with anticipated growth in usage. The baseline data collection items will remain unchanged with a few exceptions. The collection sex replaced that of gender, in compliance to Executive Order 14168. In addition, race and ethnicity will be replaced with a single combined race and/or ethnicity question that includes "Middle Eastern or North African" as a minimum category and collects detailed data beyond the minimum categories, in alignment with updated Statistical Policy Directive No. 15.

Although some follow-up collections for program experiences and tracking outcomes are not included in this renewal request, we plan to reinstate them once system usage increases to a level that allows for meaningful analyses at the program level.

A new system of records, NSF-80, Education and Training Application Data System (ETAP), (Federal Register October 2023) contain records from this system.

A. Justification

1. Necessity for the Data Collection

The U.S. National Science Foundation Education & Training Application (NSF ETAP) is a customizable, common application system designed to connect individuals—such as students and educators—with NSF-funded education and training opportunities. Initially developed for the Research Experiences for Undergraduates (REU) program, the system has since been revised and enhanced to support a diverse array of NSF programs focused on human capital development.

ETAP is designed to gather data on applicants and participants of NSF-funded opportunities; reducing burden on Principal Investigators (PIs) and applicants while enhancing NSF's ability to

monitor our education and workforce development programs and conduct future evaluations of their impact. PIs of NSF awards can utilize the ETAP system to recruit prospective participants for various NSF-funded opportunities, including research experiences, fellowships, scholarships, internships, teaching assistantships, dissertation grants, summer boot camps, and more. ETAP's centralized platform allows potential applicants to easily discover and apply for these NSF-funded programs aimed at STEM professional development and growth.

ETAP serves the future scientific community and the NSF by:
- **Providing Transparency**: Offering a centralized space for information on training opportunities, thereby reducing the burden on respondents—primarily students—who can use a single application to apply for multiple NSF-funded opportunities.

- **Improving Efficiency**: Giving new and aspiring PIs access to a robust and secure data collection system at no cost, allowing them to manage their project applications effectively, while also reducing administrative costs for established PIs.

- **Delivering Timely Data**: Offering PIs and NSF program officers immediate access to high-quality data and analytics on applicants and participants, aiding in decision-making and supporting continuous improvement efforts.

ETAP enables the NSF to gather high-quality data necessary for rigorous evidence-building activities, such as monitoring, targeted research, and evaluations of its programs. The system enhances the agency's ability to address key priorities and statutory requirements, including (1) the Foundations for Evidence-Based Policymaking Act of 2018 (Public Law 115-435), which advocates for the robust collection of data to substantiate decision-making, (2) the America COMPETES Reauthorization Act of 2010, which mandates tracking of REU program students' employment and continued enrollment in STEM fields for at least three years after their undergraduate degree (Section 514[a][6] of Public Law 111-358), and (3) the CHIPS and Science Act of 2022 (Public Law No: 117-167), which calls for new investments in STEM education while building the knowledge base on effective models for recruiting and retaining students in STEM fields.

Increasing utilization of ETAP is an NSF's performance goal aligned to 2022–2026 NSF's Strategic Plan objective 1.2 to unleash STEM talent for America, with associated metrics communicated to the public per the Government Performance and Results Act (GPRA). The ETAP system supports NSF's efforts to cultivate and expand participation in the STEM workforce, thereby advancing the development of science and technology, and enhances the agency's capacity to make well-informed program and policy decisions related to this strategic priority. By offering a secure online platform for NSF awards to attract STEM learners, ETAP collects applicant-level information directly from individuals interested in NSF-funded education and training opportunities. As ETAP use expands, the centralized online infrastructure provides NSF with more comprehensive and detailed information, including data on the characteristics of individuals participating in training activities supported by NSF programs, as well as those who apply but are not selected. This information will enable NSF to assess each program's outreach and carry out evaluations with increasing rigor.

Note that some NSF programs (or specific units within programs) mandate the use of ETAP for managing applications and collecting demographic information on participants. When ETAP usage is required, this is explicitly stated in the solicitation or award notice.

The NSF ETAP system is managed by the Evaluation and Assessment Capability Section of the NSF's Office of Integrative Activities and the Office of Chief Information Officer.

Description of the ETAP system collections (baseline data)

ETAP allows students, teachers, and other professionals to apply to multiple NSF-funded education and training opportunities, saving time and effort by submitting the same general information once through a registration form. The ETAP registration information includes:

- **Basic information** such as legal name, date of birth, and contact information that enables NSF to identify, match, and track applicants over time
- **Demographic information** and family background, such as sex, race and/or ethnicity, veteran and disability status, parental education and occupation, and Pell Grant recipient status
- **Educational background**, including degrees obtained and postsecondary enrollment status
- **Additional information** about research and work experience, and previous NSF program participation

All applicants need to complete the registration form to apply and participate in NSF-funded opportunities (describe above). Beyond the registration form, the remainder of the application varies by the opportunity applicants are applying to. PIs of NSF awards provide information about the opportunity for prospective applicants to apply to (including a description of the opportunity, eligibility requirements, program dates and location) and determine whether there any additional application requirement beyond the registration form.

- **Additional information** may include resumes, unofficial college transcripts, statement of purpose, contact information of reference writers, or answers to custom questions.

PIs and their authorized designees use ETAP to provide information needed by potential applicants, determine application requirements beyond the registration form (if any), retrieve applicant information, record admissions decisions and participation status among admitted applicants, and produce aggregate data reports.

Data collection instruments are included in the attachment.

2. Use of Information

The information collected through ETAP enables NSF to respond to NSF information needs for program monitoring and improvement, Administration priorities, and Congressional requirements, as outlined in section 1. It also enables PIs to access information they need to make admissions decisions and comply with NSF reporting requirements. User feedback ensures system enhancements are working as intended and future revisions are responsive to user needs.

3. Improved Information Technology to Reduce Burden

ETAP's centralized online infrastructure offers significant benefits for prospective applicants, the NSF Principal Investigators (PIs) aiming to recruit diverse participants, and government agencies. Key advantages of ETAP include:

- **Reduced Burden on Applicants:** Applicants can submit a single application across multiple opportunities, eliminating the need to navigate various application procedures for different NSF-funded projects outside ETAP.
- **Reduced Burden on Participants:** Demographic and educational outcomes data will be sourced from existing resources rather than relying on time-consuming surveys of former participants.
- **Support for New/Prospective PIs:** New and prospective PIs can utilize ETAP instead of creating their own application systems. This facilitates efficient dissemination of opportunity information and streamlines application management.
- **Ease for Existing PIs and Administrators:** Current PIs and administrators can shift away from maintaining individual applications. ETAP allows them to fulfill NSF reporting requirements more efficiently.
- **Benefits for the Government:** ETAP offers streamlined access to robust participant and applicant data that would otherwise entail higher burdens and lower quality.

Additionally, ETAP's functionality enhances the respondent experience and improves data quality. These features include:
- **Single Sign-On (SSO) with NSF's Account Management:** Users can access ETAP and other NSF systems, such as Research.gov, using a single set of credentials, streamlining their experience with NSF throughout their careers.
- **Data Integration with NSF Data Lake:** PIs gain real-time access to their award information, allowing for pre-population of necessary details when setting up new opportunities in ETAP.
- **Integration with National Education Data:** This feature enables applicants to easily provide educational background information (such as high school and college attended) via dropdown menus, enhancing efficiency and accuracy.

4. Efforts to Identify Duplication

The ETAP system does not duplicate efforts undertaken by the NSF, other federal agencies, or other data collection agents. ETAP was created to fill informational gaps that could not be addressed with extant data collections. A feasibility study conducted by the Science and Technology Policy Institute (STPI) (Zuckerman et al., 2016) identified the need for a new data collection system to address information gaps. ETAP was developed in response to STPI's recommendations.

ETAP leverages NSF's centralized account management profile (on Research.gov) for account authentication to avoid duplication (of efforts and of NSF accounts). Under the NSF Account Profile on Research.gov (OMB control number 3145-0269), the information that registering individuals enter to obtain an NSF ID will persist across NSF business applications and minimize duplicative efforts. Additionally, an automated pipeline process connects NSF ETAP with NSF

Data Lake for users to have access to their own proposal/award data (so that PIs do not need to provide the same information again).

5. Efforts to minimize burden on small business

Not applicable. Small businesses are not affected by this information collection.

6. Consequences of Less Frequent Data Collection

Less frequent data collection would impair NSF's ability to meet its information needs and respond effectively to Administration priorities and congressional requirements. Prior to developing ETAP, the Science and Technology Policy Institute evaluated whether NSF's existing data collections were sufficient to meet emerging legislative mandates and determined that the approach—relying on demographic reporting through standard performance reports and voluntary tracking by principal investigators—was inadequate. This assessment reinforced the need for a new, more robust data collection system.

In addition, less frequent data collection would conflict with ETAP's core purpose of connecting prospective applicants to NSF-funded opportunities, as PIs have discretion over the timing and application schedule of these opportunities within their award period—necessitating flexible, on-demand data collection

7. Special Circumstances

Not applicable.

8. Federal register announcement and consultation outside the agency

   a. Federal Register announcement

   The 60-day notice to solicit public comments was published in the Federal Register on November 22, 2024, at 89 FR 92724. No comment were received in response to the notice.

   b. Consultation outside the agency

   An expert advisory board was convened to obtain feedback on the system design and data collection when the system was first developed and pilot tested.

9. Payment or gifts to respondents

Not applicable. No incentives, payments, or gifts will be offered or given to system users.

10. Assurance of confidentiality

The ETAP system is has its own system of records, ([NSF-80 Education and Training Application Data System](#)), subject to the Privacy Act of 1974. No personal information will be released to the

public. Links to the [Privacy Act and Public Burden Statements](#) (copied below) and [Privacy Policy NSF - National Science Foundation](#) (providing access to ETAP system of record notice) appear in the website footer in all pages.

The system includes notices in three instances:

Privacy Act and Public Burden Statements *(accessible in the website footer)*

   Privacy Act Statement: The National Science Foundation (NSF) Act and other program statutes administered by NSF authorize this collection of information for purposes of one or more NSF-supported programs. NSF and its partners use this information in order to manage, track, and facilitate your participation in these programs. This information is subject to certain routine uses, including disclosure to your institution, if applicable, or others who assist or are otherwise involved in the operation or administration of these programs. These routine uses are detailed in the full system of records notice (SORN) published for Education and Training Application (ETAP) records, which is posted at [Privacy Program | NSF - U.S. National Science Foundation](#) (NSF-80). This collection of information is voluntary, but failure to provide information may affect your ability or eligibility to participate in these programs.

   Public Burden Statement: An agency may not conduct or sponsor, and a person is not required to respond to, an information collection unless it displays a valid Office of Management and Budget (OMB) control number. The OMB control number for this collection is 3145-0248. Public reporting burden for this collection of information is estimated to average 6.1 hours for applicants to apply to an opportunity, including the time for reviewing instructions. Burden estimates for principal investigators is 3.7 hours to set up opportunities and manage applications. Send comments regarding the burden estimate and any other aspect of this collection of information, including suggestions for reducing this burden, to:

   Suzanne Plimpton
   Reports Clearance Officer
   Budget and Finance Administration
   National Science Foundation
   Alexandria, VA 22314
   Please note that information provided through this system will be used for admissions decisions, audits, and research and evaluation purposes. All applicants' information will be maintained in accordance with the requirements of the Privacy Act of 1974. No personal information will be released to the public.

Rules Of Behavior *(acknowledgment required to be an ETAP user)*

   This computer system is the property of the National Science Foundation (NSF) of the Federal Government. Any system activity may be monitored and any information stored within the system may be retrieved and used by authorized personnel for law enforcement,

management, routine system operations, or other purposes. By using this computer system, you are consenting to such monitoring and information retrieval and use.

Unauthorized use of the system, including disclosure of information covered by the Privacy Act or other sensitive information, or attempts to defeat or circumvent security features, is prohibited and could result in disciplinary action, civil and/or criminal penalties. Users should be aware that they have no expectation of privacy when using the NSF-provided computer system (including any removable media used in conjunction with the system), accessing the Internet, or using electronic mail systems.

All information maintained within or retrievable through the NSF computer system, including electronic mail files, may be reviewed and retrieved by the Department of Homeland Security; NSF officials who have a legitimate reason to do so when authorized by the Director or Deputy Director; or by the Inspector General.

[Checkbox] I acknowledge the rules of behavior

Nondisclosure Agreement for Principal Investigators (and Designated) Users of this National Science Foundation Data System *(acknowledgment required to be an ETAP user with access to applicant information)*

As a principal investigator or designated user with access to this NSF Data System, I will have access to information submitted by applicants interested in participating in an NSF-sponsored program. I understand that all personally identifiable information stored in this NSF Data System will be treated as confidential. By using this NSF Data System, I also understand and agree to the following:

I will be accessing confidential information protected under the Privacy Act of 1974. Its collection and maintenance are authorized for the purposes set forth in the Privacy Act system of records notice(s) (SORN(s)) applicable to this system. See https://www.nsf.gov/privacy/. I shall use the information only for such authorized purposes, and shall not use this system to collect, maintain, or use any additional or other information about individuals that is not authorized under the applicable SORN(s).

I will be subject to laws governing confidentiality, privacy, security, and preservation of federal agency records, violation of which carries severe penalties.

My access to confidential information could be monitored.

I will access and use only the confidential information necessary to conduct work related to this NSF Data System and NSF-sponsored program at my Site.

I will not attempt to access information regarding myself, my employer, my college or university, my colleagues, collaborators, associates, or any other persons or organizations except as necessary to perform my duties in the NSF-sponsored program.

I will not disclose to any unauthorized person or organization confidential information from this NSF Data System.

After downloading personally identifiable information from this NSF Data System, I will protect this information to include controlling access only to myself or others in my organization with a business need to know, encrypting the information in transit and at rest, and destroying the information when no longer needed to perform my duties in the NSF-sponsored research program. I shall immediately report incidents involving a suspected or confirmed breach of such information to NSF IT Help Central. Such incidents include, but are not limited to, accidental disclosure of information to an unintended recipient, evidence that my system account or credentials (e.g., password) may have been compromised, etc.

Violation of these rules may result in the imposition of administrative sanctions and future ineligibility for federal grants, and may be punishable under the Privacy Act, 5 USC 552a, the Trade Secrets Act, 18 USC 1905, or other laws.
If I do not understand any of these conditions, or if I have any questions about them now or in the future, I will discontinue use of this Data System and contact the NSF Freedom of Information Act and Privacy Officer, or an assistant general counsel, in the National Science Foundation General Counsel's Office - Room 1265, (703) 292-8060.
[Checkbox] I understand that all personally identifiable information stored in this NSF system will be treated as confidential.

11. Justification for sensitive questions

Information regarding applicants' characteristics (such as race and/or ethnicity, sex and educational achievement) are collected, as this information is needed by NSF to monitor programs and respond to Administration guidance and congressional requirements, and by NSF PIs (normally included in applications used by PIs if not using ETAP). Date of birth is also collected to ensure proper identification of applicants and to obtain educational outcomes data from the NSC[1] without increasing burden on participants.

Providing demographic information is voluntary. Respondents can choose to respond, "I do not wish to provide". The system includes the following notice:

> **Providing Demographic Information is Voluntary**
> NSF seeks to increase participation in its activities among groups that are underrepresented in science and engineering. Help NSF measure progress in achieving this goal.
>
> Providing demographic information is voluntary and is not a precondition of award. On grounds of race, color, age, sex, national origin, or disability, no person shall be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving financial assistance from NSF.

12. Estimate of respondent burden

Table A1. Estimates of respondent burden by respondent type

Category of Respondent	Respondents (number)	Participation Time (hours)	Total Burden (hours)	Annual Burden (hours)
Baseline collection				
a. Principal Investigators (PIs) of an NSF award (or their designated user) providing an opportunity through ETAP	3,028	3.7	11,204	3,735
b. Applicants	117,524	6.1	716,896	238,965
c. Reference writers	166,067	0.5	83,034	27,678

---

[1] Findings from the ETAP pilot suggest that NSF can match 90% of NSF participants to NSC data—enabling access to educational outcomes with no burden to respondents—when date of birth is provided (Tuttle et al., 2024).

Totals	286,619		811,134 hours	270,378 hours

Supporting information for Table A1

a. **Principal Investigators (PIs)**

Number of respondents: We estimate that a total of 3,028 Principal Investigators (PIs) or their designees will use ETAP to manage opportunities—one respondent for each opportunity offered (e.g., participation in an REU Site in a given year

Participation Time (Burden Hours per Respondent): We estimate that PIs, or their designated users, will spend an average of 3.7 hours using the system to set up and manage applications for their NSF-funded opportunities. The actual time spent will vary depending on whether the PI chooses to run a competitive application process, the number of applications received, and the number of participants supported.

Using FY23 ETAP data to inform burden estimates, we assume:
- 75% of the opportunities will use a competitive application process
- Competitive opportunities will receive an average of 117 applicants
- Each opportunity will support an average of 9 participants
- Each NSF award will support one opportunity per year

The estimated 3.7 hours of PI burden per application cycle includes time to:

a) Provide opportunity details and submit additional application requirements
   • 0.5 hours per opportunity

b) Record admission and acceptance decisions for competitive applications
   • 2 items per applicant × 117 applicants × 1 minute per item
   • Total: 3.9 hours per opportunity

c) Record admission and acceptance decisions for noncompetitive applications
   • 2 items per participant × 9 participants × 1 minute per item
   • Total: 0.30 hours per opportunity

d) Record participation information
   • 1 item per participant × 9 participants × 1 minute per item
   • Total: 0.15 hours per opportunity

The overall burden estimate of 3.7 hours reflects a weighted average of the time required for both competitive and noncompetitive application processes

b. **Applicants**

Number of respondents: We estimate that 117,524 unique applicants will use ETAP over the course of three years to submit applications to the 3,028 NSF-funded opportunities.

The number of applicants depends on the number of opportunities offered, whether opportunities are competitive, and applicant behavior (that is, number of applications each applicant submits to ensure participation in an NSF-funded opportunity).

Based on FY23 ETAP data:
- 75% of the opportunities will be competitive (meaning 2,271 out of the estimated 3,028); 25% are noncompetitive (meaning 757 opportunities).
- Applicants apply to an average of 2.4 opportunities (if competitive).
- Each competitive opportunity receives an average of 117 applications.
- Each noncompetitive opportunity receives an average of 9 applications.

To estimate the number of unique applicants:
- Competitive: (117 applications × 2,271 opportunities) ÷ 2.4 applications per applicant = 110,711 unique applicants
- Noncompetitive: 9 applications × 757 opportunities = 6,813 applicants
- Total estimated unique applicants = 110,711 + 6,813 = 117,524

Participation time (number of burden hours per respondent): We estimate that applicants will take about 6.1 hours using ETAP to apply to training opportunities of their choice (3.25 hours to complete the common registration form, and, if applying to a competitive opportunity, an additional 3.8 hours preparing additional application requirements, weighted by the share of applicants applying to competitive and noncompetitive opportunities). These burden estimates are based on actual usage data from the precursor system (REU data system), report included in the prior information requestion (as attachment C).

c. **Reference Writers**

Number of respondents: We estimate 110,711 applicants applying to competitive opportunities (from #b above) and, therefore, 166,067 reference writers requests will be fulfilled in the system (assuming 1.5 reference letters are completed, on average).

Students applying to a competitive opportunity are asked to submit contact information for reference writers. ETAP will automatically send an email to reference letter writers inviting them to submit a letter of recommendation through ETAP. A given reference letter can be used for multiple opportunities. Based on FY23 ETAP data, on average, applicants have 1.5 completed reference letter in the system.

Participation time (number of burden hours per respondent): We estimate that it will take 0.5 hours for reference writers to complete the reference form in ETAP, which

includes a short paragraph about the applicant and requests that applicants be rated on 11 characteristics (by selecting checkboxes).

13. Cost Burden to Respondents

There are no direct costs to respondents.

14. Cost Burden to the Federal Government

The total estimated cost of this data collection is $ 4,512,604.36 which includes system operations and maintenance (including help desk support and data monitoring analytics), system development and testing for enhancements, and recruitment and communication activities. The resulting annualized cost is approximately $1,504,201.45.

15. Reason for Change in Burden

The number of respondents has increased as ETAP is expected to continue scaling its usage. Usage growth projections are based on the number of FY24 awards using the system, assuming an annual growth factor of 1.5 compared to the prior year. While the method for calculating burden remains unchanged, the burden estimates have been updated using the latest empirical evidence on respondent behavior—such as whether Principal Investigators choose to run competitive applications for their NSF-supported opportunities and the average number of opportunities each unique applicant applies to.

16. Schedule for information collection and publication

Not applicable. The system will be available to users year-round, with the exception of occasional maintenance periods or unforeseen government-related interruptions.

17. Display of OMB expiration date

The expiration date for Office of Management and Budget (OMB) approval will be displayed as shown in section 10 of this document.

18. Exception to the certification statement

There are no exceptions to the certification statement.

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

Not applicable. (No statistical methods are employed in this information collection.)

REFERENCES

Tuttle, Christina; Elizabeth Gellman, Katlyn Lee Milless, Andrés Nigenda, Silvia Robles, Cecilia Speroni, Micah Wood, and Marykate Zukiewicz. 2024. *Generating evidence for decision making with the U.S. National Science Foundation Education and Training Application (ETAP):*

Learning from a pilot of data collection efforts. Alexandria, VA: U.S. National Science Foundation.

Zuckerman, B., J. Doyle, A. Mudd, T. Jones, and G. Davis. "Assessment of the Feasibility of Tracking Participants from the National Science Foundation's Research Experiences for Undergraduates (REU) Sites Program." Final report submitted to NSF. Washington, DC: STPI, 2016.