

ATTACHMENT I – PROJECT TOPIC

Data Integration to Estimate Science, Technology, Engineering and Mathematics (STEM) Attrition and Workforce Supply: A Pilot Approach

Key Objective

The objective of this project is to improve understanding of the impact of Science, Technology, Engineering, and Mathematics (STEM) attrition that occurs within educational and workforce pathways on the supply of STEM talent required to meet future workforce demand.

For this project, STEM attrition refers to 1) enrollment choices that result in potential STEM graduates (i.e., undergraduates who declare a STEM major) moving away from STEM fields by switching majors to non-STEM fields or leaving postsecondary education before earning STEM-related degrees, licenses, or certificates and 2) those that leave the STEM workforce through the course of their career. Individuals with STEM credentials can also move between non-STEM and STEM occupations multiple times in their careers and perform different roles within the STEM workforce that are more or less related to their fields of degree.

A significant amount of research on STEM attrition has been conducted by numerous federal agencies including the National Center for Education Statistics, the National Center for Science and Engineering Statistics, the Census Bureau, and the Bureau of Labor Statistics, among others. This project will include a review of prior research to understand what has already been accomplished so that efforts are not duplicated, but instead are leveraged to enhance the outcomes of this project.

The intent is to use unique data acquisition, analytical, and data sharing approaches from federal and one (1) or more state, regional or local sources, and then create a replicable process/framework/model to continue this work across geographies in the U.S. for an informed national understanding of the impact of STEM attrition on future STEM workforce supply. This project will not propose or implement a new data collection effort to inform the study objectives. Rather, this pilot project will inform the use of existing disparate data sources, noted gaps, and innovative analytical techniques to inform the study objectives and build upon tools and services that have been implemented as part of the National Secure Data Service (NSDS) Demonstration Project to further inform a future NSDS.

Background

The [National Science Foundation's 2022 – 2026 Strategic Plan](#) states “To accelerate the advancement of discovery and learning, prepare for a world in which work is increasingly reliant upon scientific and technological skills and ensure that all citizens share in the benefits that flow from research, we must promote inclusion in the research community and STEM workforce, access to STEM learning and training and widespread STEM literacy. Our global competitiveness depends critically on the readiness of the

nation's STEM workforce, but millions of talented individuals are missing from that workforce. NSF seeks to empower these missing millions by making strategic investments in researchers and research training to harness the talents and creativity of America's diverse population. NSF also supports research into practical ways to promote a scientifically literate U.S. population that is well prepared for the economy and challenges of the 21st century."

These priorities are echoed in the [National Science Board's Vision 2030](#) which states: "Worldwide demand for STEM-capable (Science, Technology, Engineering, and Mathematics) workers keeps growing, driven by international opportunities and competition, and by rapid increases in the number of jobs that require STEM skills, including in lines of work that historically did not require S&E [Science and Engineering] knowledge. This situation will only become more urgent: by 2026, S&E jobs are predicted to grow by 13% compared with 7% growth in the overall U.S. workforce. Yet even as STEM competencies have become more essential, U.S. K-12 mathematics and science scores are well below those of many other nations and have stagnated. Women and underrepresented minorities remain inadequately represented in S&E relative to their proportions in the U.S. population. The rapid growth of S&E jobs and demographic changes have outpaced the progress that has been made in the participation of these groups in S&E."

The goals of this project are:

1. To conduct a literature review on related research that includes findings, data sources, analysis and other elements that will inform the gaps and approaches to this work.
2. To identify innovative approaches, data sources, and analytic methods to investigate the relationship between STEM attrition and workforce supply for a U.S. region or geographic area.
3. To demonstrate use of data from one or more state or local sources to inform the research topic.
4. Apply innovative data acquisition, data protection and integration techniques to produce estimates using tools developed as part of the NSDS Demonstration Project and shared services, including linking and analyzing data within the NSDS Demonstration project secure compute environment testbed.
5. To identify and compare findings between historically underrepresented groups (as specified in the NSB Vision 2030 report and other research), geographies or other areas of interest (STEM fields) that could inform policy recommendations.
6. To develop a replicable process and an analytic approach for use by researchers, policymakers, and other interested parties for analyzing data across disparate sources for different geographic locations within a future NSDS.
7. To identify lessons learned in the effort to analyze, acquire, and use disparate data sources at various levels (e.g. federal, state, local) and assess data sources fitness for use (statistical, program, etc.) to inform these research topics using pilot tools and services within an NSDS.

Information Gaps

This project will identify:

- What existing data sources (survey and administrative) are available and accessible that can currently inform this research area?
- What state or local data sources are available to inform the research area?

- What other data sources exist to answer these questions, but are not accessible? Are there alternative data sources or considerations for future analysis?
- What linking techniques are currently used and what innovative techniques could be used with accessible data sources to inform this research area?
- If additional information is needed, what can state and federal agencies do to fill in the information gaps moving forward, such as improving survey or administrative data collections?
- How can a replicable process and approach for data acquisition, data linking, and data analysis be used within a National Secure Data Service? What are the lessons learned?

Key Evidence Building Considerations

- Key focus questions (address one or more) to assess innovation in the following areas: data acquisition, data security, data linking, data quality (fitness for use), privacy, and engagement:
 - What are key challenges with locating, acquiring, accessing, linking, and using disparate STEM data and information to inform STEM attrition and the future workforce?
 - Which novel techniques for data, privacy and confidentiality protections can be used?
 - What estimation methods, including the possibility of model-based methods, are best suited to fill information gaps? Are the resulting data of an appropriate quality level to inform policy discussions?
 - Are the resulting data and models fit to use to inform policy discussions and to make data available more equitably?
 - What approaches can be taken to broaden access to the resulting data while upholding confidentiality and privacy requirements?
 - What types of collaboration and stakeholder engagement is needed to help inform these questions, data, and analysis?

Deliverables

At a minimum, offerors will provide the following if selected for an award. Additional deliverables may be required.

- Monthly status reports on progress towards project objectives.
- Quarterly lessons learned from ongoing tasks and activities.
- A literature review of STEM attrition and workforce supply as well as the ongoing work in this area to include noted gaps in data, analysis, or other dimensions.
- An analytical plan that includes a recommended approach for data acquisition, including state and local sources, analysis, and techniques/methods that addresses this project's research questions of interest by geography and underrepresented groups.
- A report of overall findings, including estimates that inform the role of STEM attrition on future workforce supply. The report will also describe the data, tools and methods used in this work, to also include risk assessments and lessons learned for activities conducted within a secure compute environment.
- A report that highlights study findings with policy-relevant groups, geographies, or other factors that may be used by policy stakeholders.
- A framework or model for replicating the study's processes and approaches to be used by other researchers across different geographies.