

MEMORANDUM

Date: April 22, 2016

Revised May 17, 2016

To: Shelly Wilkie Martinez, Desk Officer
Office of Management and Budget

From: John Gawalt, Director
National Center for Science and Engineering Statistics

Via: Suzanne Plimpton, Clearance Officer
National Science Foundation (NSF)

Subject: Notification of data collection under generic clearance

The purpose of this memorandum is to inform you of NSF's plan to collect data on users' needs under the generic clearance for survey improvement projects (OMB #3145-0174). This study is part of a larger set of activities assessing how well our current dissemination tools and data products meet the needs of our users. We have completed a user analysis and have proposed a study for usability testing of NCSES tools under a separate generic clearance. This request describes our planned data collection for an analysis of NCSES data user needs.

Background

The National Science Foundation's (NSF) National Center for Science and Engineering Statistics (NCSES) is the principal source of analytical and statistical reports, data, and related information that describe and provide insight into the nation's science and engineering resources. All of the Center's data are released in electronic format on NSF's centrally administered web server (<http://www.nsf.gov/statistics/>). NCSES also manages additional external and internal servers that supplement the central server's content.

With increasing technological advancements and the changing needs and preferences of its stakeholders, NCSES strives to make improvements to its data dissemination activities by identifying those processes, products and tools that may require reassessment and refinement. NCSES faces the same challenges that other federal statistical agencies experience in understanding and meeting the needs of its data users. We participate in and benefit from knowledge and experiences shared during discussions among members of the Federal Interagency Dissemination Group, which brings together the data dissemination leaders within the federal statistical system. The group is charged with identifying the ongoing challenges in dissemination and sharing of best practices, including usability testing, to help support the missions of their respective agencies.

The goal of this project is to obtain insights on users' needs related to science and engineering statistical data by conducting a small, exploratory and focused data collection. The results of the study will guide NCSES in prioritizing resources for

refining NCSES data tools and products. Results also will inform NCSES decisions on the future development of new products that better the needs of NCSES’s target audience.

Recruitment

NCSES plans to send a short, web-based data collection instrument to gather information about user needs. (See Attachment A for the survey instrument.) Potential participants will be selected to be broadly representative of a participant group and to have knowledge and/or interest in science and engineering statistics to represent the current user or potential user base, as opposed to the general population. We plan to recruit participants from the U.S. through targeted email without any regard to geography. (See Attachment A for contact scripts.) To maximize response rates as well as reduce the burden of filtering out potential participants that do not have an interest in science and engineering statistics, participants will be targeted based on their past interactions with information on the science and engineering enterprise. Sources for recruitment include individuals who have contacted an NSF survey manager for assistance, signed up for the National Center for Science and Engineering RSS feed, or published papers or reports that incorporate data on the science and engineering enterprise. We expect to field the instrument in the spring of 2016 and continue for 4 weeks.

Our goal is to have a uniform distribution of responses from the groups listed in the first column of the table below. Invitee names will be obtained from the sources listed in the second column. Names will be chosen from the lists to obtain a diverse pool based on institutional affiliation.

Participant Group	Source of participants
Policy Analysts	<ul style="list-style-type: none"> • Contact lists from survey managers of analysts of government contractors • Authors of papers or other publications that use NCSES data or other science and engineering (S&E) data
Media	<ul style="list-style-type: none"> • Contact lists from survey managers of press requests.
Academia	<ul style="list-style-type: none"> • Restricted use license holders • Authors of papers or other publications that use NCSES data or other S&E data • Contact lists from survey managers of university-based institutional researchers interested in academic benchmarking
Industry	<ul style="list-style-type: none"> • Contact list from survey managers of people with industry affiliations
Nonprofit Organizations	<ul style="list-style-type: none"> • Contact lists from survey managers of people with nonprofit affiliations • Authors of papers or other publications that use NCSES data or other S&E data
Casual Information Seekers	<ul style="list-style-type: none"> • People registered for the RSS feed (random sample of non-.edu, non-.gov, and non-foreign addresses) • <u>Contact lists from program managers of people who appear to be casual information seekers (i.e., they do not fall into the above categories).</u>

We plan to recruit participants through email invitations. An invitation will be sent out first to 120 people (20 per group) with a goal of obtaining 60 total responses (10 per group). (Contact scripts are included in Attachment A.) We expect that the response rates will vary drastically between groups; for example, academics that use NCSES data probably will be very likely to respond, while casual information seekers probably will be less likely to respond. Since this is an exploratory activity, there is no prior work from which to infer exact response rates. After two weeks, up to an additional 280 people may be invited from participant groups that did not attain the goal of 10 responses. Exact number will be based on response rates from the first wave.

No personal or sensitive information will be collected. Since this is a targeted survey, the names of respondents will be known. The survey software will keep track if an invitee has responded or not so that reminders will only be sent to those who did not respond. Invitee names and respondent names will be kept confidential (e.g., invitees' and respondents' names will not be published.) However, ~~The software will be set to record if a token has been used but will~~ not save the token as part of the response record, making the responses anonymous.

Burden Information

We expect to invite a maximum of 400 people with the goal of obtaining responses from 10 users per participant group. We expect the recruiting process to take an average of five minutes per invitee, resulting in approximately 33.3 hours of burden (400 people x 5 minutes \approx 33.3 hours). The estimated time for completion for the instrument is 15 minutes. If every invitee responds, the total maximum burden for the data collection activity would be 100 hours (400 responses x 15 minutes = 100 hours). Thus, we estimate a total burden of approximately 133 hours for this research.

Incentive Payments

There are no incentive payments.

Contact Information

The contact person for questions regarding this data collection is:

May Aydin
Supervisory Program Director
National Science Foundation
(703) 292-4977
maydin@nsf.gov

Attachment
A – Contact script and data collection instrument

cc: Joydip Kundu

May Aydin
Rebecca L. Morrison
Tyler Fox