MEMORANDUM

Date: December 20, 2019

To: Margo Schwab, Desk Officer
             Office of Management and Budget

From: Emilda B. Rivers, Director
             National Center for Science and Engineering Statistics (NCSES)
             National Science Foundation (NSF)

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 NCSES'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 the current website, dissemination tools, and data products meet the needs of users.

Background

The National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF) is the principal source of analytical and statistical reports, data, and related publications that describe and provide insight into the nation's science and engineering resources. The Center's publications and related content are released in electronic format on NCSES's various website URLs (https://ncses.nsf.gov/; https://nsf.gov/statistics; https://ncsesdata.nsf.gov/).

NCSES will redesign its website and data dissemination systems to support content under a unified subdomain URL (https://ncses.nsf.gov), supported by a new content management system (CMS) where possible. The current site organizes information based on the name of the survey or publication series from which the information was collected. This may or may not be the best way to do it, since users may think of the information differently, in the context of their particular applications.

In preparation for a complete redesign of the main NCSES website and integrated data system (i.e., interactive table tool and metadata explorer), NCSES is looking to identify intuitive, high-level NCSES topic and subtopic metadata labels that will serve as the navigational gateway to exploring NCSES data and information. The revised topics and subtopics will be an entry point to navigate end users to the information or data of interest.

This project involves two studies. The purpose of the first study is to query users and potential users about their needs as they explore science and engineering data and information (called the "interview study"). The second study, called the "tree testing study," will test a set of topics and subtopics developed per an understanding of users'

needs. Tree testing is a usability technique for evaluating the findability of topics in a website. The results of both studies will be used to develop an information architecture for the new website. A follow-on activity (for which another generic clearance for data collection will be submitted) will develop the user interface for the new website. Both studies, in sequence, are required to redesign the NCSES website.

Recruitment

For both studies, NCSES plans to recruit participants that represent specific groups of data users (see table below). They will not be representative of the general population. NCSES plans to recruit participants from the U.S. through targeted email without any regard to geography. (See Attachment A for the initial recruiting e-mail.) 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 NCSES survey manager for assistance, signed up for the National Center for Science and Engineering Statistics (NCSES) email updates subscription, or published papers or reports that incorporate data on the science and engineering enterprise. Names will be chosen from the lists to obtain a diverse pool based on institutional affiliation.

In addition to email invitations, NCSES plans to recruit participants for the interview study through posts on social media and a banner on the current NCSES website with a goal of obtaining 36 total interviewees. (Recruiting messages are included in Attachments A and B.) An invitation will be sent to 21,230 people, assuming that only about 1% of recipients will open, read, and respond to the invitation, and that about 20% of those who do will be eligible to participate, per the demographics being selected for (See Attachment C for Screener). NCSES expects that the response rates will vary drastically between groups; for example, academics that use NCSES data 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. Recruitment will occur in waves to limit the potential burden on respondents. We will solicit participation from approximate 21,130 people currently subscribed to the NCSES GovDelivery email listserv, the number fluctuates periodically due to new subscribers and unsubscribes from the list. We will also solicit participation from an additional estimate of 100 people, using a list maintained by NCSES staff of individuals who are familiar with the information that NCSES maintains.

The tree testing study, which will occur several weeks after the interviews, will collect data from up to 100 participants. NCSES plans to invite the 36 interviewees to participate in tree testing, as well as the larger set of respondents who completed the screener. If necessary, a third wave of solicitation will be conducted with 5,000 people to find participants for tree testing.

Table 1: Number and Types of Participants

Participant Group	Source of Participants	Characteristics of Participant type	Target # of Participants for interviews	Target # of Participants for tree testing
Government Analysts	· Social media groups · 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 · Social media groups · Visitors to the current website	· Users employed by a Statistical Agency (e.g., Bureau of the Census, Bureau of Labor Statistics), other, Government Agency, or internal NSF e.g., survey owners · Users use data to prepare papers, publications, or analyses	8	22
Media	· Contact lists from survey managers of press requests · Social media groups · Visitors to the current website	· User prepares documents and uses data to support journalism or writing for public consumption (e.g., the topic is education is a path to citizenship and will use data on Science and Engineering in the US)	8	22
Academia	· 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 · Social media groups · Visitors to the current website	· User is employed by a university or is a student and searches for data to support academic use · Authors prepares papers, publications or analyses that use NCSES or other S&E data	8	22
Nonprofit Organizations	· 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	· User is employed by a non-profit agency · Authors prepare papers, publications or analyses that use NCSES or other S&E data	6	16
Industry	· Contact list from survey managers of people with industry affiliations · Social media groups · Visitors to the current website	· User is employed by a for-profit, non-government, non-statistical/policy agency, non-NSF · Users use data to prepare papers, publications, or analyses	4	12
Casual Information Seekers	· People registered for the NCSES email updates subscription (random sample of non-.edu, non-.gov, and non-foreign addresses) · Contact lists from program	· User may not be applying S&E data in a professional capacity · User may or may not be doing analysis of data, but is using summaries and descriptions to make decisions	2	6

Study Description

This project involves two studies: the first involves interviews, the second involves tree testing. The first study will use an exploratory and open-ended approach to query users and potential users about their needs as they explore science and engineering data and information. Up to thirty-six semi-structured interviews of 60 minutes will be conducted over a web-based video-conferencing platform. During the interviews, conducted via video conference, participants will be asked to share their screen to show us how they search for information on the NCSES site (See Attachment F). They will be asked to talk aloud as they conduct their searches to reveal how they search, what terms they use, and what challenges they encounter. Participants will also be asked to perform searches for provided examples and talk aloud as they go through a task they may be unfamiliar with. The interview will elicit search terms, strategies and preferences.

An NCSES contractor will be conducting the interviews and NCSES staff will be observing. Participants will be told that their participation is voluntary, and they will be informed of the OMB control number (3145-0174; see Attachments F and I). NCSES expects to field the instrument in early 2020 and continue for 4 weeks. No sensitive information will be collected.

For the second study, research participants will be asked to use a draft topic/sub-topic taxonomy to navigate to the place they would expect to complete a task (see Attachments G and H). The tasks will be developed based on input from the research interviews previously conducted. The test will be conducted using tree testing, a usability technique for evaluating how well the topics work in terms of findability. It is also known as 'reverse card sorting' or 'card-based classification'. Tree testing is done on a simplified text version of the topic structure – without the influence of navigation aids and visual design. This will help articulate where and why people can't find what they are looking for. The test will help answer the questions:
- Do the topic labels make sense to people?
- Do the content groupings under topics make sense to people?
- Can people find the information they want easily and quickly? If not, what's stopping them?

The topic and subtopic structure will be tested in two rounds, each with up to 100 participants. The topics/subtopics will be revised between rounds based on input from the first round. Participants in the second round may be drawn from the first-round participant pool, as those interviewed would be more likely to participate in the tree testing.

The tests are unmoderated and will be conducted over an automated web-based testing platform. The tool, made by Optimal Workshop, functions similarly to a dynamic, web-based survey tool; participants click through alternate text-based topic hierarchies (i.e. "trees") to try to navigate to find items in different categories. By quantifying the ease of finding items, NCSES can measure the effectiveness of alternative topical hierarchies.

Participants' inputs during the tests will be recorded, and an NCSES contractor will analyze and synthesize the data. The analysis will answer questions such as:

- Could users successfully find particular items in the tree?
- Could they find those items directly, without having to backtrack?
- If they couldn't find items, where did they go astray?
- Could they choose between topics quickly, without having to think too much?
- Overall, which parts of the tree worked well, and which fell down?

Burden Information

For first study, NCSES expects to invite a maximum of 21,230 people with the goal of obtaining responses from up to 36 users (see Attachments A, B, C, D, and E). NCSES assumes that all of those solicited will read the invitation (21,230 * 1 minute = 354 hours), but only 1% of those solicited (212 people) will respond to the invitation by completing the screener (212 * 5 minutes = 18 hours). NCSES expects the recruiting process to result in approximately 372 hours of burden (354 hours + 18 hours = 372 hours).

The estimated time for completion for the 36 interviews is 75 minutes, including 15 minutes for scheduling (36 * 75 minutes = 45 hours). Thus, the total burden estimate for the first study is 417 hours.

For the second study (tree testing), NCSES expects to recruit up to 100 participants. NCSES will invite those who completed interviews in the first study, and those who completed the screener but were not selected for an interview. If 100 tree tests are not completed as a result of those invitations, a maximum of 5,000 additional people will be invited. NCSES assumes that all of those solicited will read the invitation (requiring 1 minute each). NCSES expects this process to result in up to approximately 87 hours of recruiting burden (5,200 people * 1 minute = 87 hours).

The estimated time for completion for the tree testing data collection is 15 minutes. If the up to 100 participants who are selected complete the tree testing, the total maximum burden for the data collection activity would be ~25 hours (100 responses * 15 minutes = 25 hours). Thus, the estimate is a total burden of approximately 112 hours (87 recruiting hours + 25 testing hours) for tree testing.

The maximum total burden for both interviews and tree testing is estimated to be 529 hours.

Incentive Payments

A $40 Amazon gift card will be sent as an honorarium upon completion of the interview. A $25 Amazon gift card will be sent as an honorarium upon completion of the tree test. These represent tokens of appreciation for the participants' time and effort, not

considered payment for their time as professionals, and provides value and expresses thanks (see Attachment J).

Contact Information

The contact person for questions regarding this data collection is:

May Aydin
Supervisory Program Director
National Center for Science and Engineering Statistics
National Science Foundation
(703) 292-4977
maydin@nsf.gov

Attachment A – Initial recruiting e-mail
Attachment B – Initial recruiting post for social media and homepage
Attachment C – Screener
Attachment D – Scheduling email for eligible participants
Attachment E – Thank you email for non-eligible participants
Attachment F – User needs assessment interview guide
Attachment G – Tree-testing recruiting e-mail
Attachment H – Protocol for tree testing
Attachment I – Interview consent form
Attachment J – Thank you email with honorarium (electronic or mail)

cc: May Aydin
     Rebecca L. Morrison

Attachment A: Initial recruiting e-mail

Dear <Full Name>,

You are invited to participate in a feedback session. It would take about an hour and you would be given a $40 gift card as a thank you. We have identified you as having an interest in science and engineering (S&E) data, either because you have contacted us or have published on S&E topics.

In order to meet your current and future needs, we at the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation would like to better understand how you use our data products. NCSES is the principal source of analytical and statistical reports, data, and related information that describe and provide insight into the nation's S&E resources. The Center distributes a variety of data on S&E topics including research and development (R&D) funding, R&D performance, and the S&E workforce.

If you'd like to talk with us for approximately an hour about your needs related to science and engineering data, please **click the link below** to answer a few questions that determine your eligibility for the study. This should take less than five minutes.

<link to screener survey>

If you meet our criteria for participation, Amy Goldmacher, from our research partner, The Understanding Group (TUG), will be in touch with you to schedule an interview. At the conclusion of the interview, as a thank you, we will send you a $40 Amazon Gift Card. If we do not call you for an interview, we do hope that you will be interested in participating in other research to improve our services.

Your answers will be kept confidential. Your feedback is a critical element to our ability to learn and improve. Your participation is voluntary. This study is authorized by law (42 U.S.C. 1862 Section 3.a.6.). The OMB control number for this study is 3145-0174.

If you have questions, please do not hesitate to contact me. We look forward to your participation.

Sincerely,
May Aydin
<signature, contact information>

Attachment B: Initial recruiting post for social media and homepage

Text for Social Media

Do you use science and engineering data in your work? If so, we'd like to hear from you! Please click the link below and take 5 minutes to see if you meet participation criteria. If you do, we will interview you and send you a [AMOUNT] Amazon Gift Card as a thank you.

[Link]

Text for website home page

Join the NCSES Web User Community ListServ.

NCSES is always looking for feedback from our data tool and web users. Your feedback helps NCSES to continue to improve the functionality of our online tools.

Sign up for the NCSES listserv and get information quickly and directly to and from NCSES via email. To subscribe, please send an email to listserv@listserv.nsf.gov with the text of the message as follows:

Subscribe [listname] [subscriber's name]
Example: subscribe NCSES-WEB-USER-COMMUNITY John Smith

If you no longer wish to be included on the distribution list, you can elect to be removed from the list at any time. Instructions for unsubscribing will be included at the end of each list message.

Attachment C: Screener

Thanks for being willing to participate in our study of needs related to science and engineering data! This study is sponsored by the National Center for Science and Engineering Statistics within the National Science Foundation (https://www.ncses.nsf.gov/). Please answer the following questions - this should take less than 5 minutes.

1. Which best describes the industry you work in?
   a. Statistical Agency (e.g., Bureau of the Census, Bureau of Labor Statistics)
   b. Other Government Agency
   c. Policy Agency
   d. Non-Profit
   e. Academia
   f. Media
   g. For-Profit Organization (other than media)
   h. Other: please describe

2. About how often do you search for science and engineering data or reports?
   a. Daily
   b. Weekly
   c. Monthly
   d. Quarterly
   e. Yearly
   f. Less than Yearly

3. Have you ever used NCSES data or reports?
   a. Yes
   b. No
   c. Don't know

4. Which of the following other sources of science and engineering data or reports have you or do you used on a regular basis? Check all that apply.
   a. OECD
   b. US Census Bureau
   c. Data.gov
   d. The World Bank
   e. UNESCO
   f. Other: please describe

5. How do you primarily use the science and engineering data or reports you search for? Check all that apply:
   a. I gather summarized data for someone else's use
   b. I prepare a publication that requires data to support conclusions or decisions

     c.   I find and use raw science and engineering data to come up with my own analyses or statistics

     d.   I find evidence to answer a question

     e.   Other: please describe

6. Where are you located (city, state)?

7. What is your first name and last name?

8. Are you employed? If so, what is your employer name or URL?

9. What is your preferred email address?

Attachment D: Scheduling email for eligible participants

Participants who go through screener and who meet participation criteria are sent the following email inviting them to schedule via Calendly (all reminders and confirmations handled automatically through Calendly).

Hi <First Name>,

Thank you for being willing to participate in our research about your needs related to science and engineering data on behalf of the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation (NSF).

We would like to obtain more feedback from you during a remote interview that would be held at your convenience. The interview would last no longer than 60 minutes, and no preparation is necessary. Please click the link to find a time that works best for your schedule.

<insert link to calendly scheduler>

We will meet via Zoom, a web-based video-conferencing tool. You will be asked to share your computer screen during the interview.

If none of the available times work for you, please let me know and we'll find a time.

Thanks again,

Amy

<Contact information>

Attachment E: Thank you email for non-eligible participants

Participants who go through screener and are not selected for participation in either the interview or the tree testing are sent the following as thanks for going through the screener.

Hi <Name>,

Thank you for taking the time to answer a few questions from the National Center for Science and Engineering Statistics (NCSES) regarding science and engineering data. Though you do not meet the eligibility criteria for this particular study at this time, we hope that you will be interested in participating in future research projects.

Sincerely,

May Aydin
<signature, contact information>

Attachment F: User needs assessment interview guide

Thanks for joining us today. We are with The Understanding Group, an information architecture agency working with the National Center for Science and Engineering Statistics (NCSES) on part of a website redesign project. We are looking for your perspective on how you search for and find content you need. We want to understand how you use the site in order to make recommendations for improvement. We've got 60 minutes with you. We'll have a discussion and we'll ask you to show us how you actually use the site.

There are no right or wrong answers - we're gathering perspectives and will aggregate responses.

Is it OK to record this session for note-taking purposes only?

[If respondent says yes, then walk through the consent form, answer any questions, then turn on the recorder, and get verbal consent. If respondent says no, then walk through the consent form, noting that no recording will be made, but that participation is voluntary.]

Before starting, please review this consent form. Do you have any questions? May I have your verbal consent to conduct the interview?

Introduction

1. Tell us a little about yourself. What is your role in your organization? How long have you been there?

2. What's unique about the science and engineering data you look for and how you look for these data?

Conducting a Science and Engineering Search

Let's talk about how you search for science and engineering data.

3. Tell us about the most recent time you needed science and engineering data: What were you looking for?

4. What prompted the search?

5. Where (what site) did you go to first? Can you show us? (Pass screen sharing control over)

6. What did you expect to get when you searched? What was helpful (probe for: topics, organization, content, discoverability?) in getting to what you needed?

7. Can you recall an instance of stumbling into or serendipitously finding valuable sources of science and engineering information or data?

Search on NCSES

Let's go through the same process on the NCSES site [navigate to NCSES for known item finding]. Please don't share anything confidential, but please let's go through the search you told us about.

8. How would you do the search you told us about?

9. What do you need to see when you search? (topics, subtopics, discoverability?)

10. Tell us about how you would search for something you were interested in on the NCSES website?

11. Is that process different from what you would do when you are not sure what you're looking for? Please explain.

12. When in your searches is the "topics" navigation element important? (probe for nestedness)

13. Do you search by Sector? (e.g., Healthcare Industry)

    a. Why/why not?

14. Tell us about what you expect when you see "search by Sector"?

15. Do you search by Demographics? (e.g., specific variables like African American Men)

16. Tell us about what you expect when you see "search by Demographics"?

17. Given that this website has both Sector and Demographic content types, which of those do you look for and why?

18. Could these content types be organized better?

    a. How and why/why not?

    b. Does your preference ever change? Depending on what?

19. Are there any other challenges when searching on the site?

Search on NCSES for new topic

Now we'd like to have you do a search on an example subject (or a couple, time permitting). I'll give you an example, and I'd like you to talk aloud about how you would search for this as you find it.

Examples:

A

Find: "an analysis comparing the number of scientific publications in 2017 in the US with that of China."

Title: *Science and Engineering Publication Output Trends: 2017 Shows U.S. Output Level Slightly Below That of China but the United States Maintains Lead with Highly Cited Publications*

Link: https://www.nsf.gov/statistics/2019/nsf19317/

B

Find: "a comparison of the US research and development budget in 2016 and 2017."

Title: *U.S. R&D Increased by $22 Billion in 2016, to $515 Billion; Estimates for 2017 Indicate a Rise to $542 Billion*

Link: https://www.nsf.gov/statistics/2019/nsf19308/

C

Find: "results from a survey of research and development in colleges and universities from 2018."

Title: *Higher Education Research and Development Survey, FY2018*

Link: https://www.nsf.gov/statistics/srvyherd/#tabs-2   or
https://ncsesdata.nsf.gov/herd/2018/

D

Find: "data on the number of bachelor's degrees awarded to women vs men across different fields in 2016."

Title: *Table 5-2: Bachelor's degrees awarded, by field and sex: 2006–16*

Link: https://ncses.nsf.gov/pubs/nsf19304/data   or
https://ncses.nsf.gov/pubs/nsf19304/assets/data/tables/wmpd19-sr-tab05-002.xlsx

20. How would you start to search for science and engineering data for this? Then what would you do?

21. What are you looking for when you search?

22. What do you need to see (content) to help you find these data?

Wrap up

23. What would make you return to NCSES site?

24. If you could wave a magic wand, how would the NCSES site be organized to work for you?

25. Is there anything that would you like to add that we didn't ask about?

26. Where should we send your honorarium (electronic or mail - confirm email or get mailing address per preference)?

27. Would you be willing to be contacted again to participate in future studies or provide feedback to NCSES?

28.

Attachment G: Tree-testing recruiting e-mail

Dear <Full Name>,

You are invited to participate in a short feedback session. It would take 15 minutes and you would be given a $25 gift card as a thank you. We have identified you as having an interest in science and engineering (S&E) data, either because you have contacted us or have published on S&E topics.

In order to meet your current and future needs, we at the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation would like to better understand how you navigate our site to find our data products. NCSES is the principal source of analytical and statistical reports, data, and related information that describe and provide insight into the nation's S&E resources. The Center distributes a variety of data on S&E topics including research and development (R&D) funding, R&D performance, and the S&E workforce.

If you have 15 minutes to test new ways of organizing topics on our website, please click the link below.

<link to tree test>

At the conclusion of the test, as a thank you, we will send you a $25 Amazon Gift Card. We do hope that you will be interested in participating in other research to improve our services.

Your answers will be kept confidential. Your feedback is a critical element to our ability to learn and improve. Your participation is voluntary. This study is authorized by law (42 U.S.C. 1862 Section 3.a.6.). The OMB control number for this study is 3145-0174.

If you have questions, please do not hesitate to contact me. We look forward to your participation.

Sincerely,
May Aydin
<signature, contact information>

Attachment H: Protocol for tree testing

Note: the following is a DRAFT outline of the participant experience for tree testing. A final protocol with exact questions will depend upon information gathered from the interviews conducted in the first study.

Steps in the Participant experience:

- Participants will sign-in to web page.

- They will be presented with a consent form and asked to confirm that they have read it and click "approve" in order to continue.

- They will be asked questions in survey format and then presented with a "tree" of topics/subtopics.

- They will be given tasks to complete in the context of that tree such as "using the topics, please select where you would go if you were looking for [XYZ]".

- The participant can then navigate through the tree to find what they think is the answer and move on to the next question.

The instrument will track their activities as they move through the tree. We will analyze how successful our topics were at guiding people to the right place. We will then revise the topic structure and re-run the test.

Attachment I: Interview consent form

National Science Foundation

NCSES Website Usability Testing
Consent Form

CONSENT FORM

The National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF) designs, conducts, and sponsors surveys on science, engineering, and technology. In order to produce the best information possible, the National Science Foundation routinely evaluates its data tools and products.

You have volunteered to take part in a study to improve one or more of the National Center for Science and Engineering Statistics (NCSES) website and data products. In order to have a complete record of your comments, your interview session will be audio taped (and/or videotaped). We plan to use the tapes to improve the tools and products. Staff directly involved in this research project will have access to the tapes. In addition, a transcription service may have access to the audiotapes, to provide us with a written version of your comments. The tapes may also be used for training others to conduct this type of research, and in presentations to professional audiences. Specific results or quotes will not be attributed directly to you in any notes, summaries, or reports.

This study is authorized by law (42 U.S.C. 1862 Section 3.a.6.). The OMB control number for this study is 3145-0174.

I have volunteered to participate in this study, and I give permission for my tapes to be used for the purposes stated above.

_____   _____
Researcher's Signature   Participant's Signature

_____   _____
Printed Name   Printed Name

_____   _____
Date   Date

2415 Eisenhower Avenue | Alexandria, VA 22314

Attachment J: Thank you email with honorarium (electronic or mail)

Hi <Name>,

Thank you so much for participating in our research about your needs related to science and engineering data. Your feedback is a critical element to the National Center for Science and Engineering Statistics' (NCSES's) ability to learn and improve.

<Attached or included> is your honorarium. We hope that you will be interested in participating in future research projects to improve our services.

Thanks again,

Amy

<Contact information>