

# Survey of Researchers at U.S. Research Universities on Software Use, Development, and Training

## I. Introduction

Research software is essential and/or central to a large fraction of today's research projects, and many others are dependent on software. Currently, we do not have a formal study to provide data on this. This study will benefit the subjects and society by providing data on this topic. In the UK, a similar study was done two years ago, which led to a national institute to promote understanding of the role of research software and to provide both training and recognition to its developers.

The objective of the research is to assess software use and development among researchers at U.S. universities, and training associated with such software development. We also want to explore differences in software use, development, and training between specific groups among academic faculty and staff. This involves examining how the degree of research activity of the university, academic discipline, gender, and professional title correlate to software use, development, and training for faculty and staff. The significance of this research is that the results contribute to assessing the culture of software use in academia. Furthermore, the results might be used to evaluate the need for funding in the area of software development and training in all academic departments.

## II. Literature Review

A similar study was conducted in the United Kingdom. The results of this study headlined that:

- 92% of academics use research software
- 69% say that their research would not be practical without it
- 56% develop their own software (worryingly, 21% of those have no training in software development)
- 70% of male researchers develop their own software, and only 30% of female researchers do so

These figures are specific to U.K. Universities. Specifically, 15 Russell Group universities were polled. The results were based on the responses of 417 randomly selected researchers at these institutions.

One of the limitations of this study was that while it uncovered facts regarding software use, it did not examine the causes of these disparities.

### **III. Methodology**

Since we aim to understand the culture of software in academia, we have selected 15 U.S. universities that conduct research and have doctoral programs. A sample of 5 universities will be selected from 3 of the Carnegie research grouping (Doctoral Universities: Highest Research Activity, Doctoral Universities: Higher Research Activity, and Doctoral Universities: Moderate Research Activity). These rankings are meant to be similar to that of the Russell group universities. Universities where we have prior connection will be given priority in selection.

In order to recruit subjects for our study, we plan to contact the Vice Chancellor of Research (or equivalent role) by e-mail at each of these institutions and ask them to forward our survey to their researchers for them to complete. (Exactly which researchers they have a means of targeting will vary from university to university, but will include faculty and staff in all cases, and potentially postdocs and graduate students in some cases). Completion of the survey will be completely voluntary. Researchers will not be significantly associated with the subjects. The institutions that will provide access to the subjects will be the 15 universities that will be contacted (See attached list of universities). Subjects are not being selected from records. The solicitation will be entirely through direct e-mail with the subjects and through the institutional “gatekeepers” of the administrators at each university.

Equitable selection of subjects is assured through our use of the VCRs (or equivalent) at the universities to distribute the survey to all individuals that they determine to be involved in research by their institution. As there is no selection after acquiring the contact information, there is no disparity in selection between groups in our subject pool. The only criterion for inclusion of subjects is that they are considered researchers by their institution. There are no anticipated measures and protocols that will be used to screen applicants. The inclusion/exclusion criteria will be assessed by the Vice Chancellors of Research (or equivalent roles) at each institution based on their knowledge of who currently conducts research at their universities.

The participants will fill out a web survey conducted through a Google Form (see attached document). The questions included on the survey are multiple choice and short response. The survey will take approximately 30 minutes to complete and will be sent to the subjects in August 2016. The activity will take place online and can be completed at anytime within a month of sending the survey with a tentative deadline of September 1st 2016.

Data will be collected through the submission of the survey by the participants. The responses will be stored on a Google Drive belonging to the Responsible Principal Investigator. This will allow the investigators to view and interpret the responses to the survey questions. Since the survey is not asking for any identifying information of the participants, the investigators will not have the ability to trace responses to subject identities.

The data will be stored on Google in the system that the builds up the survey answers, and the computer of the Responsible Principal Investigator. Only

investigators working on this project will have access to the data. No identifying information on participants will be stored.

#### **IV. Recommendations**

There are not yet specific plans for disseminating the research. We expect to use the research results in a number of ways.

We will provide a summary of responses to each Vice Chancellor of Research or equivalent, showing how their university researchers answered compared to researchers from all the universities we survey, and compared to researchers from universities with the same Carnegie research classification.

We will write a blog or magazine article about the results, and present them at one or more workshops and conferences, including in a conference paper.

We will use the results in a proposal to NSF and/or private foundations, to improve the state of the field as discovered in the survey.

#### **V. References**

Hettrick, S. (2014, December 4). It's impossible to conduct research without software, say 7 out of 10 UK researchers. Retrieved July 7, 2016, from <http://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers>

## Universities Included in Study (Tentative)

### **Doctoral Universities: Highest Research Activity**

1. University of Illinois at Urbana-Champaign (Public)
2. University of Iowa (Public)
3. Pennsylvania State University (Public)
4. Tulane University (Private)
5. George Washington University (Private)
6. University of Mississippi (Public)
7. Washington University in St. Louis (Private)

### **Doctoral Universities: Higher Research Activity**

1. Illinois Institute of Technology (Private)
2. Indiana University–Purdue University Indianapolis (Public)
3. Michigan Technological University (Public)
4. Portland State University (Public)
5. Rensselaer Polytechnic Institute (Private)
6. Southern Illinois University (Public)
7. University of California at Merced (Public)

### **Doctoral Universities: Moderate Research Activity**

1. California State University, Fullerton (Public)
2. DePaul University (Private)
3. Louisiana Tech University (Public)
4. Rochester Institute of Technology (Private)
5. Mayo Graduate School (Private)