

## **Interview Protocol for XALT Data Reuse**

IRB: 2015-06-0022

### **Procedures**

Interview participants will be interviewed using a set of open ended questions related to XALT data reuse. The participant may be asked additional questions during the interview depending on their answers to planned questions. A sample dataset will be reviewed by the participant during the interview.

### **Location**

Interviews will occur face-to-face when possible. Those participants who cannot meet face-to-face will be interviewed via phone, email, or web-based video conferencing.

### **Measures**

The study measures include interviews designed specifically for this project.

### **Participants**

Target population: professionals working in High Performance Computing administration and software development and use, data analyst interested in big data usage, and social scientists that have expressed interest in the XALT data to understand scientific computing trends, and data curators.

### **Inclusion/Exclusion**

Interviewees will be selected based on their role indicated in the target population. Participants must have an active role in the scientific research community.

### **Risks**

There are no foreseen risks in the proposed study.

## **Recruitment**

Recruitment will be done via invitation.

## **Privacy and Confidentiality**

Participants will remain anonymous for their participation in the interview process.

## **Obtaining Informed Consent**

Participants will be given a description of the study prior to their interview via email. The description will also give the participant an opportunity to remain anonymous for the purposes of publication.

**Interview** The interview will take approximately 30 minutes.

## **STUDY**

### **Introduction**

The interviewee will provide an overview of the XALT data, its provenance and scope and content. The question and answer discussion involves showing the dataset to the interviewee.

### **Questions**

What is your professional/academic background?

#### *Data Availability*

Where would you expect to find this data?

Do you currently have access to similar data?

#### *Data Use*

How might you use this data in your work?

What information elements in the dataset are the most useful to your inquiry?

#### *Data Format*

During analysis, what span of time would you need to work with?  
(month, quarter, year)

What format would you prefer the data to be distributed?

#### *Data Analysis*

What platform would you use to analyze the data?

How will you download the data? (high speed internet connection, personal computer, HPC resource)